

724

MODALITY AND THE THEORY OF MEANING

[AN EXAMINATION OF THE PROGRAMMES
OF DAVIDSON, DUMMETT AND MONTAGUE]

by

RICHARD MERVYN SPENCER-SMITH

Submitted as a dissertation for the degree of Ph.D.
in the University of London.

Bedford College, October 1981

ProQuest Number: 10098472

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10098472

Published by ProQuest LLC(2016). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ABSTRACT

R.M. Spencer-Smith: MODALITY AND THE THEORY OF MEANING

Each of the three main chapters of this thesis is concerned with a different style of theorising about the semantics of natural language and in particular with the way they would tackle expressions of modality. What unites the three approaches is their commitment to the study of a language through a systematic theory which will account for all its sentences, according to some general principle such as: meaning = truth conditions. They diverge widely on the implementation of this idea, ie. on the aims and form of a theory of meaning. The first promises a very spartan kind of theory; there is therefore considerable interest in discovering how such an austere method will manage to handle the intricacies of intensionality. To this end several ways of coping are examined in the first chapter. The second approach permits itself a much richer means of describing the semantics of a language. Consequently, the concern of the second chapter is not so much with coping, as with marshalling these more powerful resources into a detailed analysis of some of the linguistic manifestations of modality. The third approach is, as yet, more often critical than constructive. It seeks to replace theories of the first two kinds, founded on what it sees as an unjustified realist metaphysics, with a more cognitive semantics. In the course of the thesis, different manifestations of modal concepts within sentences of natural language are examined, ranging from the 'outermost', sentential operator occurrences to the 'innermost' occurrences where the modality is interwoven into the property expressed by a simple predicate. Thus in the last chapter, the import of the criticisms raised by the third approach is assessed with special reference to dispositional predicates.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	3
<u>CHAPTER 1</u> DAVIDSON AND TRUTH THEORY.....	4
1 Introduction.....	4
2 Truth-Theoretic Preliminaries and the Wallace Argument.....	32
3 The Paratactic Theory.....	40
4 Operators in the ML.....	55
5 Modality as Quantification.....	71
<u>CHAPTER 2</u> MONTAGUE AND MODEL THEORY.....	80
1 Structural Semantics and Model Theory.....	80
2 Modality and Montague Grammar.....	99
3 Modality: Meaning.....	113
4 The Problem.....	134
<u>CHAPTER 3</u> DUMMETT AND ANTI-REALIST THEORY.....	150
1 Introduction.....	150
2 Dispositions.....	159
3 Possible Worlds and Anti-Realist Modality..	184
4 Semantics and Metaphysics.....	200
CONCLUSION.....	219
SOME NOTATION.....	229
FOOTNOTES.....	231
BIBLIOGRAPHY.....	237

ACKNOWLEDGEMENTS

I am deeply indebted to my tutor Hans Kamp; without his forbearance and his criticism this thesis would not have been possible. For sundry other encouragement, guidance and tutelage I would especially like to thank Bill Hart, David Wiggins and Andrew Wright.

1: INTRODUCTION

It is important to stress at the outset that Donald Davidson approaches the topic of meaning by aiming to illuminate the nature of interpretation - more specifically, the interpretation of the utterances of speakers of some natural language. Interpretation of utterances is the linguistic department of the more general project of interpreting behaviour - of making sense of people, as McDowell puts it. In Davidson's conception of the accomplishment of the task relating to linguistic behaviour, the interpreter will be engaged in the construction of a theory, a process in which two inputs figure centrally. The first input will consist of that which is to be interpreted: descriptions of utterances which the theory must transform into specifications of their content (and force - though I shall largely ignore this component of meaning in what follows). Any OL sentence ϕ , suitably designated, as input must be paired with a sentence p of the ML which gives its content(cp.[67], p.44). The second input will consist in the evidence upon which an interpretative theory can be non-question-beggingly founded. The evidence will consist in observation of the circumstances in which speakers of the language produce various utterances. A speaker's utterance of a sentence is taken as prima facie evidence that he holds it true - that he believes whatever it is that it expresses(cp.[20], p.322). The empirical business of theory construction is to be constrained by the principle of charity - that overall the

pairings $\varphi...p$ devised by the theory enable the interpreter to ascribe an intelligible set of beliefs and desires to the speakers of the language under study (cp. [13], p. 313). This brief résumé will suffice, I hope, as a background for discussion of Davidson's claim that the central component of an interpretative theory of meaning for a language will take the form of a Tarskian theory of truth for it.

In framing his famous adequacy condition on definitions of truth, Convention T, Tarski laid it down that a definition would be acceptable if it was, in effect, an interpretative theory - if it paired every sentence φ of the OL with a translation p of it into the ML. If it was to define a predicate T , with pretensions to being truth, the pairings should naturally be of the form $T(\varphi, L) \equiv p$. If someone had a theory meeting Tarski's requirement they could interpret utterances of sentences from the OL. But as yet there would be no philosophical illumination of meaning: Convention T assumes that the work is already done, that φ means that p . Davidson's approach promises illumination by showing how we can earn ourselves the right to see the Tarskian pairings $T(\varphi, L) \equiv p$ as interpretative, and thus to read them as: φ means in L that p (cp. [22], p. 37). The extensional form of the Tarskian pairings promises an ideally non-question-begging framework for the project, since one cannot simply reverse Tarski's conceptual priorities and get meaning out of truth.

(1) $T(\text{Snow is white, English}) \equiv \text{grass is green}$

is true, but hardly interpretative. It is appealing to think that by subjecting candidate theories to the kind of empirical testing indicated above, "the constraining need to match truth with truth throughout the language"¹ will achieve the

desired effect of eliminating the true but non-meaning-stating theories. But there is a drawback: merely leaving the constraints to filter out the bad theories will leave you with a semantic theory (or theories, depending on indeterminacy) cast in the Tarskian mould which, because it is extensional, will have no defence against parasitical deformations of its axioms which render it uninterpretative (cp.[40],p.xvii). In order to gain immunity we must return to Tarski's original approach - reimpose the constraint that the pairings $\varphi...p$ are interpretative, the difference being that this time it is explained how, not assumed that, this is achieved.

By the disquotational effect of the truth predicate, we know that if we write out the pairings in the $T(\varphi, L) \equiv p$ form, we can regard the predicate T as "True". But why should we fill the gap this way, rather than using an intensional biconditional or, since our interest is not especially in defining a truth predicate, with "means that" - or even "means the same as"? I shall consider the last, translational format, in a little while. As to "means that", it is unclear how we could use it in a general and systematic way. How, for example, could we state an axiom for and using it? And even for predicates, an axiom

$(\forall \alpha)(\alpha \text{ is brave means that } R(\alpha, \text{English}) \text{ is brave})$ is rather dubious. If I say Jones is brave, I don't mean that the reference, in English, of Jones is brave, I mean Jones is brave. The same might be said of the intensional biconditional - that it is only at the final stage that we get properly meaning stating pairings. Another reason, I think, relates to the holistic nature of the empirical investigation Davidson envisages interpretation to be. It

is with the interpretative pairings that the theory is closest to the evidence, since the data concern the holding true of whole sentences. What the theory is putting forward for testing, in its T-sentences, are hypotheses about extensional equivalence - the evidence available to the radical interpreter relates not to intensions, but merely to extensions. The requirement of interpretativeness is imposed on the theory as a whole, "that the totality of T-sentences should.... optimally fit evidence about sentences held true by native speakers"([20],p.326). There is no point tampering with the T-sentences, though of course, if the theory is adequate the pairings will be counterfactual supporting, will be translational(cp.[22],p.36).

We shall return to the question of the form of the pairings later. For the moment, let us note the special vulnerability which Davidson's espousal of holism lends to the project. Should the Tarskian truth-theoretic format be incapable of accomodating a certain natural construction, it is not open to the interpreter, one who aims "to describe and understand" a language, to say "So much the worse for that idiom". If that should turn out to be the case, since for the holist there is no giving the meaning of one sentence of a language without giving it for them all, the failure would in principle vitiate the semantics proposed for sentences not employing the recalcitrant idiom (cp.[13], p.308 and [106],pp.233-4).

When the radical interpreter is banished to the jungle he will be well schooled in domestic science: he will be acquainted with all the means available to a Tarskian truth theorist for treating the various kinds of construction that occur within his own language. His construction

of a theory for the alien tongue will be constrained not only by such requirements as Charity, but also by certain formal principles, and it is to these I now turn. Not any old theory of truth will serve the ends of an interpreting theory of meaning, and the constraints are designed to impose the desired form on candidate theories. They are as follows (from [17], p.19). A) Scrutability: "the theory [should] provide a method for deciding, given an arbitrary sentence, what its meaning is." If the idea of a theory which would systematically give the meanings of all the sentences of a language can in principle be realised, this seems a natural requirement for it to meet, and I shall not discuss it further.

B) Learnability, or finitude: "An acceptable theory should ... account for the meaning (or conditions of truth) of every sentence by analysing it as composed, in truth-relevant ways, of elements drawn from a finite stock." At some level, a semantic theory for a natural language must operate on a logically disciplined (though not regimented) version of that language. How much logical form one perceives in a sentence depends initially on one's interests. Bardot is a good actress may come out as $\varphi(b)$, or even just an unstructured ψ , for the purposes of some arguments. Neither of these forms would be satisfactory for a semantic theory of our language. The policy of welding good and actress into a single predicate φ would require an infinity of axioms if it were to be generally implemented, since good can modify a potential infinity of phrases. And if that, then why not simply an infinite list, one axiom for each sentence in the language? Such a theory would give the meanings, but it would not display the semantic structure

which must exist if the language is to be learnable by finite creatures. Demanding learnability requires of a theory that it discern at least enough structure to enable the list of semantical clauses for the expressions it treats as unstructured to be of finite length. Similarly, the number of semantical clauses which deal with the compounding of expressions must be finite. A minimal imposition of the constraint would apply directly to the number of basis and recursion clauses of the theory - a reasonable demand on any kind of semantic theory - without going into questions about speaker's acquisition and mastery of the language. Of course, any theory could get by with a special axiom for good actress treated indivisibly, and thus cheat when interpreting Bardot is a good actress. But if it is to do justice to all the other occurrences of good and actress, it would somewhere have to have the means to tell us what those familiar words are doing in that sentence. Such an axiom would be dispensable in theory, but helpful in practice. In practice, simplicity dictates that we take short cuts, suppressing a parameter here, welding a phrase up there, when our interest is taken by some other construction - on the assumption that the longer route is available in principle.

C) Homophony, or Same Concepts: "the statements of the truth conditions for individual sentences entailed by the theory should ... draw upon the same concepts as the sentences whose truth conditions they state." According to Wallace ([106], p.226), a homophonic theory is one a) whose ML includes its OL and b) which moreover matches each closed sentence of the OL (mentioned) with itself (used) in the $T(\phi, L) \equiv p$ biconditionals. One might wish to separate these out as two concepts of homophony, freeing the requirement of

sentence by sentence sameness of concept - b), in effect - for foreign as well as domestic interpretation. C) seems to me to be the most problematic of the three constraints. There are at least two ways of understanding it. The first way, consistent with its status as a formal constraint on candidate theories of interpretation, is the "natural" interpretation Davidson gives it when introducing it. This is as a general constraint, imposed before interpretation gets started, eliminating any theory whose form dictates that it employ semantic concepts in the statement of the truth conditions of any sentence, whatever its subject matter (cp.[23],p.248). This seems unexceptionable as long as it is not understood as a prohibition on other departments of semantics (cp.[19],p.79 - it deems them "irrelevant to present purposes", ie. to interpretation). For example, when we have reached a certain level of competence, our acquisition of new names may proceed via beliefs that are semantical in nature. When someone starts to use a freshly encountered name, Jones say, the only belief about Jones he may have is that Jones is whoever Jones names. Since all English speakers have the ability to pick up new names, it would therefore be incautious to rule that a full description of speakers' competence should not avail itself of semantic concepts. But as I have already suggested, when it comes to interpretation if someone says merely that Jones did such and such, whatever their beliefs about Jones, it would misrepresent the nature of their utterance to use some such phrase as "the bearer of Jones" in the interpretative pairing.

A second way of understanding C) would be as precisely the requirement of interpretativeness. The constraint

appeals to the very notion - sameness of concept - which we expect an account of interpretation to furnish us with. What could be more sensible than to require, when interpreting someone's utterance, rather than for example analysing their words, that you should aim to get as close to their words with your own as you can. This is the task of "effecting a full reckoning in essentially the same terms as speakers themselves employ of the semantic and compositional resources of the language", in Wiggins' words ([11], p.337). Just as B) demands that a truth theory read enough (structure) into sentences, so C) demands that it not read too much (structure or anything else) into them. But to impose it at this stage may suggest that it is superfluous because already in force, or else question-begging, if taken to require of the interpreter that he have insight into whether a concept figuring on the LHS of a T-sentence is the same as one used on the RHS. Even if radically interpreting English in English, we cannot assume that ([13], p.312). But there is another facet of the Davidsonian project to which it is relevant: schooling the radical interpreter in domestic science, ie. doing truth theory for English in English, relying unashamedly on one's native insight into the language in order to see how best to handle the various parts of speech. The thinking up of ways in which to process different locutions is not an empirical matter in the strong sense; rather it is the construction of the framework in which empirical interpretation can take place. The idea is to have plenty of equipment for the jungle, so that when the interpreter hypothesises that a certain alien noise is adjectival, say, he will know the form of clause it requires. Understood in this context, C) is nothing other than a sharp

formulation of Tarski's requirement (though cp.[61],p.348).

In order to fully assess the force of the constraint, it is necessary to examine the different stages in the project of interpretation. This can be looked at in two ways. Firstly, consider how a finished theory would semantically process a foreign sentence (or, if all interpretation is 'work in progress', a theory which is finished enough to be generating interpretations). This could be either a two-stage or a three-stage affair. Consider the following simplified specimen²:

(2a) Um Mitternacht schlug Shem Shaun

↑

I

↓

b) $(\exists e)(\text{schlug}(\text{Shem}, \text{Shaun}, e) \ \& \ \text{um}(\text{Mitternacht}, e))$

↑

II

↓

c) $(\exists e)(\text{hit}(\text{Shem}, \text{Shaun}, e) \ \& \ \text{at}(\text{midnight}, e))$

↑

III

↓

d) At midnight Shem hit Shaun

I: OL syntax - transformations which eg. insert/delete variables. II: truth-theoretic interpretation ___ is true iff... via eg. the machinery of sequences. III: ML syntax. Finally, between 2a) and either 2c) or 2d) is induced the relation ___ is true iff..., hopefully readable as ___ means that....

The second way of looking at it is to consider the number of stages in the construction of the semantic component of an interpretative theory. For Davidson there are two: a direct definition of truth for the disciplined sentences of the language, the deep structures (II), and a systematic matching of sentences in surface dress to those for which truth has been directly defined ([20],p.320). Thus in interpreting an OL sentence involving adverbial

modification of an action sentence (2a), we might end up with a disciplined ML sentence containing explicit quantification over events (2c). Seen in this light the following passage of Davidson's suggests that constraint C) is operative at stage II, requiring the theorist to tidy away any bits of ontology, such as sequences, which his theory needs in order to derive T-sentences:

the striking thing about T-sentences is that whatever ontological wheels must turn, in the end a T-sentence states the truth conditions of a sentence using resources no richer than, because the same as, those of the sentence itself. Unless the original sentence mentions possible worlds, intensional entities, properties or propositions, the statement of its truth conditions does not. ([20], p.319)

So constraint C) would appear to be rather weak, merely ensuring that eg. quantification over events is preserved in the transition from OL to ML. By contrast, what principally constrains stage I is the suitability of the deep structures revealed by the transformation rules for handling by the truth theory: principle B), and extensionality. Davidson acknowledges that the relations at stage I may be very complicated (eg. in [21], p.319).

This (pairing of 2a) with 2c)) may be all very well for a Davidsonian theory mark one, but it hardly seems to fit the bill for theories which are to be interpretative in the sense of that word cultivated under the revised Davidsonian format. The problem is that by imposing the interpretativeness constraint on the whole theory, even if the best philosophical arguments show it to be necessary that: 2c) if and only if 2d), still the divergence in explicit structure and ontology will be undesirable. One seemingly drastic remedy to this state of affairs would be to impose constraint C) as a condition on the first stage.

This may seem unattractive in that the deep paraphrases are bound to diverge to some extent in their state of discipline, and moreover the divergence was precisely that which enabled certain surface forms to be handled truth-theoretically. However, there is nothing wrong in requiring that C) be met as far as possible at stage I - that ceteris paribus, if a theory has surface and disciplined forms which are closer than those of another it is to be preferred. For, as we shall shortly see, even at stage II C) cannot be enforced absolutely and without exception. The idea is, then, that we should come to semantic investigation with a preference for homophonic theories, as Evans puts it (in [38], p.271). Divergence, such as that between 2a) and 2c), is to be tolerated only when unavoidable.

A strategy different to that explicitly adopted by Davidson, though perhaps complementary to it, would involve the completion of the third stage III, where possible, ending up at 2d) rather than 2c), for example. It only seems fair that, in interpreting a sentence, we are allowed to reverse, in the ML, the discipline imposed in the OL. In the process of interpretation, the theorist may uncover hidden deep structure and ontological commitment of a sentence, though these will not show up in the final pairing. Since interpretation requires discovering first order quantificational structure in a language (cp. [20], p.323), for each language there will be transformation rules linking the surface with first order structures. In which case the interpreter should be reasonably confident that for any deep structure which turns up in the ML at stage II, he will be able to convert it into a more colloquial form in that, his language. We should still require that homophony be

preserved, as far as possible, at every stage³. For example, if the OL contains different sentences All men are mortal and Every man is mortal, whose form is undistinguished by the $(\forall v)$ notation, it would be preferable to paraphrase them using (all v) and (every v) respectively, if the precise flavour of these sentences is to be recaptured in translation. It should be noted that this three-stage approach, whether it acknowledges it explicitly or not, must involve sentences in the ML being mentioned as well as used.

However, there still remains an important discrepancy between the truth conditional approach and constraint C). It concerns the interpretation of contextually sensitive material. Our ordinary notion of truth is unperturbed by such material - we can say for example It is true that that is yours, because we are using this notion of truth within a particular context of utterance. Using this notion as a predicate of sentences we could perfectly acceptably say That is yours is true if and only if that is yours, because the correctness of an ascription of truth is here itself a matter which will change from context to context (cp.[102],p.128). But the perspective of the semantic theorist is acontextual - he must state the truth conditions of a sentence, whatever the utterer, time etc. of the context might be. And if he is to do this, he cannot use words which are like those in the sentence in being contextually sensitive. If, to continue the example, he was really using that, the thing he pointed to would enter into the truth conditions of any sentence using that demonstrative. So, as some have complained (eg.[61],p.349) and as Davidson is fully aware, truth theory for a language with indexicals is going to require the employment of

semantical concepts, such as that of being the thing demonstrated by an utterer u at a time t .

Let us consider the case of ambiguity in a little more detail. The occasion where we have difficulty in understanding an ambiguous sentence is the exception rather than the rule - somehow context prods us in the right direction. There are two kinds of case to consider, and Davidson has proposed two kinds of strategy to deal with them. In the first, the theorist can match ambiguity with ambiguity. For this Davidson has suggested a homophonic approach (at p.316 of [13]). Thus we might have:

- (3) $(\forall u)(\forall t)(T_c(\text{Schmidt ist komisch}, u, t, \text{German}) \equiv \text{Schmidt is funny at } t)$.

Gupta has assumed that a similar line would be taken on structural ambiguity; his example being

- (4) $T(\text{The president of U.S. was a Democrat}, \text{English}) \equiv \text{the president of U.S. was a Democrat.}$

Gupta comments on this

as a sentence of English [4]] is ambiguous and it has two readings. And the semantic theory can be understood as asserting [4]] on just one of these readings or on both these readings. If the former then the theory does not take into account the ambiguity of 'the president of U.S. was a Democrat'. If the latter then at least one reading of [4]] is false (in December 1976). In either case the semantic theory that entails [4]] is unacceptable. ([48], p.443)

In the case where 4) is false, Gupta presumably holds that the RHS is read as false, but on the LHS an ambiguous sentence of English is unambiguously denoted, and truth truly ascribed to it. But surely one cannot simply ascribe truth to an ambiguous sentence type, taken out of context (unless, perhaps, it is true on all its readings. But ex hypothesi this is not the case here). 3) does not make this mistake, employing the utterer/time relativised truth

prédicate. But it does have an anomolous feature. At a first glance, it looks as if the theorist is cleverly exploiting the match in ambiguity of the word/sentence meanings to state the truth conditions of Schmidt ist komisch. But how is he using "Schmidt is funny" to do this? Not in the ordinary way, intending it in one of the two senses (since he must do justice to both), nor yet as a pun, in both senses (since many utterers of the designated German sentence would not use it this way), nor vaguely or disjunctively, in neither sense in particular. Suppose, for example, that someone says John is at the bank, where in the context it is quite generally understood to mean a particular financial institution. If John is not at the bank, in this sense, then even if he happened to be at the (most salient) river bank, the utterance would be counted as false. Once an interpretation has been selected, its truth relative to a different reading is simply irrelevant. In aiming to cover all possible utterances of the sentence, the theorist covers none, since he cannot be committed to any one particular use (of the interpreting sentence). Although the pairing he comes up with seems admirable, it is not clear that the theorist can state any correct or coherent condition with 3).

Since structural ambiguity is likely to be handled by mapping the surface form onto two disambiguated deep structures, and since the approach in 3) cannot anyway generalise to the case where the ML cannot capture the ambiguities in the OL, let us turn to Davidson's other proposal concerning lexical ambiguity ([17], p.20). As an illustration:

(5) $(\forall u)(\forall t)(T_c(\text{Schmidt ist } \overset{b}{\text{böse}}, u, t, \text{German}) \equiv (\text{either$

Schmidt is angry at t and the circumstances surrounding u at t meet condition c, or Schmidt is evil at t and the circumstances surrounding u at t meet condition c'))

Since 5) is formulated generally, for the interpretation of the sentence uttered by any u at any time t, the conditions c and c' should correspond in some way to the general ability of German hearers to move from an utterance of the specimen sentence to one or other interpretation. It is difficult to see what specific conditions could realistically accomplish this task, if not something fairly abstract along the lines Kripke suggests, involving the semantical notion of utterer's meaning ([61], p.349). For example, the first disjunct would read: Schmidt is angry at t and bose means, in u's mouth at t, the same as angry (or: ...and u meant, at t, that Schmidt is angry).

Even if one can execute the strategy of 5) within a formal theory (which K.P. Parsons doubts, in [76], p.389), it has the obvious disadvantage of violating the 'same concepts' requirement. Dropping the qualifications for c and c' in 5) would lead to a T-sentence which, although true, would not be meaning-stating. Not only does Schmidt ist b^ose not have the structure of a disjunction, it doesn't mean that: either Schmidt is angry or he is evil. In a given context, it usually either means one, or it means the other. An austere treatment of ambiguity should therefore aim at disjoining T-sentences, thus:

- (6) $(\forall u)(\forall t)((T_c(\text{Schmidt ist b^ose}, u, t, \text{German}) \equiv \text{Schmidt is angry at t}) \text{ or } (T_c(\text{Schmidt ist b^ose}, u, t, \text{German}) \equiv \text{Schmidt is evil at t})).$

But this only highlights the need for the homophony destroying constraints c and c', for 6) allows that

Schmidt ist böse is true if, say, Schmidt is evil, regardless of the sense it was intended in or understood as having when uttered. (There is an interesting discrepancy here. We can say that, for any utterer *u* and time *t*, either Schmidt ist böse means, in *u*'s mouth at *t*, that Schmidt is angry at *t* or it means, in *u*'s mouth at *t*, that Schmidt is evil at *t*. The pairing is right but the truth-theoretic wrapping is wrong.)

It seems to me that the best way for a truth theorist to handle lexical ambiguity is to change tack and treat it more along the lines already suggested for structural ambiguity. At the level of deep structure there is not to be one lexical item böse, but two: böse₁ and böse₂ - likewise for komisch. A lexically ambiguous German sentence is then treated as having two (or more) disciplined versions (eg. Böse₁ (Schmidt) and Böse₂ (Schmidt)), each with its own distinct truth conditions. An utterance of Schmidt ist böse by a German speaker is therefore treated as in fact one of two possible utterances.

We have seen something of the shape of the deliverances of a Davidsonian theory of meaning. Let us now compare how a translation manual fares as an interpretative theory. The radical translator is subject to the same empirical constraints as the radical interpreter; exactly the same amount of evidence is available to both of them. Moreover, we can adapt the demands of scrutability, finitude and homophony to suit the format of the translation manual. The translator is interested in defining a function TR from expressions of one language to those of another - his own, I shall assume in general. The translation manual which involves three languages: translated, or source (*L*₁),

translator's (L_2) and translating, or target (L_3), although it exhibits the translational format in its full generality, is from the radical point of view something derivative, obtained from the two language case. One would first have to produce a manual from L_1 to L_2 in L_2 , then translate that to: L_1 to L_3 in L_2 , or take some foreign translator's manual, from L_1 to L_3 in L_3 , and convert that.

A manual for German to English should aim at defining $TR(\varphi, G) = \psi$, to be read: the translation of φ from German is ψ , or more roughly: in German, φ means (the same as) ψ .

Consider how the requirement of homophony fares. When ordinarily translating from German, it would be quite natural to utter the following kind of thing:

(7) 'Jener Hund gehört mir' means 'That dog belongs to me'.

You wouldn't hear the quotation marks, but they must be present. This is so because 7) does not require any accompanying act of demonstration, and quotation has the effect of sealing a demonstrative off from its immediate context of utterance (cp. §1.3 below). Unlike the case of contextual truth considered earlier, no demonstrative is being used (at least, not in the standard manner). In his clauses for singular terms, we can therefore expect the translator to incorporate both

(8) $TR(\text{Schmidt}, G) = \text{Schmidt}$, and

(9) $TR(\text{jener}, G) = \text{that}$.

As well as other rules for basic expressions, eg.:

(10) $(\forall \varphi)(TR(\text{Mensch}\varphi, G) = \text{man}TR(\varphi, G))$

(11) $(\forall \varphi)(TR(\text{möglicherweise}\varphi, G) = \text{possibly}TR(\varphi, G))$

the translator will have to treat of structure. One way would be to provide rules which transform syntactic rules of L_1 into syntactic rules of L_3 . But we may also conceive of

translation as operating at stage II of the interpretative process considered above. If interpretation at least requires the imposition or revelation of quantificational structure in the foreign tongue, the requirement applies with equal justification to translation⁴. What reason could there be for thinking that translation manuals need be any different at stage I, where OL forms are disciplined?

The matching of type with type may give the impression that translation is merely a syntactic operation, much like the transformations of OL syntax, say. After all, both kinds of transformation must preserve meaning, as far as possible. The difference is that the radical translator's TR function embodies a complex empirical hypothesis about what the expressions of L_1 mean. A translation manual does not make the futile attempt to break out of the circle of language; on the contrary, it exploits the understanding an English speaker has of his own language in conveying to him the meanings of the foreign expressions. 7), or a formal version of it served up by manual, enables an English speaker to understand any German utterance of the German sentence by supplying him with a sentence which is its nearest equivalent in his language. He can, as it were, substitute a pattern of noise which he does understand, in the mouth of a speaker, for one he does not. Because the translator is not using the words (in their normal senses) as the interpreter is, he is not obliged to eliminate the contextual variability by recourse to concepts not explicitly expressed by the foreign sentence. He can aim to pair a foreign sentence with one which matches its dependence on context in exactly the same way - translation is, where possible, character-preserving (in Kaplan's terminology).

He thus produces pairings, similar to 4), which comply better with the requirement of homophony than those of the interpreter. Consider the case of ambiguity once more. When the translator cannot match what he supposes to be an ambiguity in L_1 with a word in L_3 , we can imagine him producing a dual clause, such as:

$$(12) (\forall \varphi)(\text{TR}(\text{"böse"}\varphi, G) = \text{evilTR}(\varphi, G) \text{ and } \text{TR}(\text{"böse"}\varphi, G) = \text{angryTR}(\varphi, G)).$$

The translation of a sentence containing böse will bifurcate, and serve up two sentences (it means this:..., and it means this:....), presenting to the English speaker explicitly as a choice what will hopefully be the same choice as is faced tacitly by the German speaker. When he can find a word which matches the ambiguity of the foreign word, as with komisch and funny, the translator can do what the truth theorist could not: treat it in a single clause.

Whether a translation manual can count as a theory of meaning depends, as we shall see, on the nature of quotation, as it is employed in radical translation. More obviously, it depends on what precisely the requirements are on a theory of meaning. For example, if it is simply required to be, in Wallace's words, "a theory such that anyone who knows it is in a position to understand every sentence of the language," ([108], p.45; cp. [24], p.131) one might think from what has just been said that a translation manual will suffice. But it is often held - as Wallace goes on to argue - that the form of pairings $\varphi \dots p$ which a translation manual delivers makes it unsuitable to be deemed a theory of meaning. The essential point is that a truth-theoretic theory of meaning for German in English states facts which relate German sentences to conditions in the

world, not to sentences in another language. To be clear about the argument, it should be emphasised that from the point of view of radically understanding German in English, both a truth theory and a translation manual could summarise, in their different ways, the same empirical data about the meanings of the foreign expressions. The complaint against translation manuals is the parochial nature of the facts they state: that whereas one can understand Es regnet by knowing, in the appropriate way, that it is true in German if and only if it is raining, without knowing English in particular, such understanding is not imparted to one ignorant of English but equipped with the knowledge that Es regnet means in German the same as It is raining. A French-only speaker who was informed as follows:

- (13) L'énoncé allemand Es regnet est traduit (en anglais) comme It is raining,
would be in such a position.

What fact does this state:

- (14) The German sentence Es regnet translates (into English) as It is raining ?

We determine this in part by deciding how we would translate this sentence into another language - French, for example. On the view critical of translation manuals, 14) goes over into 13), but an alternative possibility would be:

- (15) L'énoncé allemand Es regnet est traduit (en français) comme Il pleut.

It is here that one's views on the nature of quotation become crucial. For, while we must admit that a French-only speaker could understand 13) without knowing what Es regnet means, the point would be irrelevant to the aspirations of a translation manual which delivers 14), if the fact it states

thereby would be stated in French by means of 15).

Two things are striking about the proposed translation of 14) into 15): one quotation (but not the other) has been translated, and "English" has become "français". Why should we think that some quotations are translatable? The short answer to this is: it happens. In translating from a German novel, "Er hat 'Es regnet' gesagt" would become "He said 'It is raining'". An actual example, from more technical literature: Tarski's T-sentence ",es schneit" ist eine wahre Aussage dann und nur dann, wenn es schneit" has been translated "'it is snowing' is a true sentence if and only if it is snowing" ([99],p.269; [100],p.156). We may suspect that there is a principle at work here which sorts translatable occurrences of quotation from nontranslatable ones, and which has something to do with whether an interpretation is presupposed for the quoted material. But even so, what of the shift from English to French in the language referred to in the proposed translation? Surely "Preserve reference" is an inviolable principle of translation practice. But consider Hart's example:

(16) This sentence is in English.

This may go over into German as

(17) Dieser Satz ist auf englisch,

preserving self-reference, but not truth-value, or

(18) 16) ist auf englisch,

preserving reference but not self-reference, or

(19) Dieser Satz ist auf deutsch,

preserving self-reference and truth-value, but changing the reference to both language and sentence. Which of 17)-19)

is the most apposite rendition of 16) will depend on the context in which that sentence is produced, especially on the

point and purpose of its utterance or inscription (cp.[8], p.142) - ie. whether reference or self-reference is more important to it. The example serves to illustrate Hart's principle that translation preserves reference if and only if it does not preserve self-reference. Since one cannot preserve them both, the general directive to preserve reference must be qualified to allow cases of self-reference to constitute an exception.

The potential significance of Hart's principle becomes even more apparent when we recall that on an autonymous view of quotation, a quotation itself may constitute a case of self-reference - a special use of words to refer to themselves. The need for an account which discerns semantic structure in quotations is evident, as Davidson has pointed out, from our ability to process a potential infinity of them, in that at least we know what the reference of any of them is. On an autonymous view, quotation marks do not perform a semantically significant function; they serve merely to disambiguate, by signalling the presence of the special, autonymous use of the words. In principle, they are dispensable. The semantically significant units are the words concerned, and concatenation. One thing the autonymous theory has going for it is that in spoken language quotation marks are dispensed with (except when quote, unquote and so on are used, to emphasise reportage of the precise words uttered). To take Geach's example (from [43],p.81), if we hear the two sentences Man is an animal and 'Man' is a noun we know from the content of the latter that the word man is being used to refer to itself, even though nothing signals this explicitly. We can even kill two birds with one stone, using and mentioning,

as is evident from Quine's well known example

(20) Giorgionne was so-called because of his size.

Giorgionne is being used in two senses: primarily, to refer to the man to whom, in the assertion of the sentence, a certain property is ascribed; and also to refer to itself, in order for that ascription to be intelligible.

The following argument is invalid:

(21) 'Flann O'Brien' is a nom de plume

(22) Flann O'Brien is Brian O'Nolan , so

(23) 'Brian O'Nolan' is a nom de plume.

In "Refernce and Modality" Quine draws this conclusion:

"The principle of substitutivity should not be extended to contexts in which the name to be supplanted occurs without referring simply to the object" ([89], pp.139-40). The autonymous theorist agrees with Quine on this, but does not concur with him in concluding that a word is not a genuine constituent of a quotation in which it occurs - as little as "corn" is in "unicorn". His whole point is that it occurs, but is used to refer to itself, not its standard reference (if it has one). One phenomenon which the theory cannot deal with is the mention of expressions which are not words - as in McDowell's example: The machine printed out 'zxwt prt gjh'. It is integral to the theory to distinguish quotation-as-autonymy and the 'mere mention' of expressions, for which we need some account in terms of pattern (or sound) mimicry. Ex hypothesi, the structure we have to discern in such cases must be finer than units of a word's length - the letters, numerals or other symbols which comprise the mentioned expression. But if we must have this more inclusive account, why bother with the autonymous theory at all? The answer to this is that - if we have not

seen reason enough already - a spelling account is so implausible for the ordinary, central cases of quotation (cp. [69], pp. 215-16).

An important question for the autonomous theorist, and especially relevant in the context of our general discussion, is: what account is to be given of the quotation in one language of some foreign expression or sentence? In answering this, it will help to have some perspective on the identity conditions of words. There are two naturally competing views of word (type) identity. On one view, words are language-specific entities, whereas semantical relations are not language relative. We should not talk of the truth-in-English of a sentence, but of an English sentence's being true. On the other view, these relativities are reversed (recall the sketch of Tarski above). Both views can cite ordinary locutions in their favour. For the former, we have eg. "the English word..." etc. (cp. 14) above). For the latter, we might say: "The very same word, 'Mist', means something totally different in German". Although there is perhaps no firm fact of the matter in dispute here, some considerations seem to me to favour the latter view. For one thing, it can hardly be denied that the identity of symbol tokens, be they inscriptions or even just flags or flag-wavings, cannot be tied to one particular language. It is uncontroversial that the same pattern of sounds can occur in more than one language. As Davidson says, "the sounds 'Empedokles liebt' do fairly well as a German or an English sentence, in one case saying that Empedokles loved and in the other telling us what he did from the top of Etna" ([15], p. 146). Secondly, as McDowell has pointed out ([69], p. 221), it is odd to make the identity of a word turn on the identity

of a language, given the problematic nature of the latter. From an intuitive point of view, a natural language is a great amorphous thing, with very frayed edges; from a Quinean point of view, its individuation is as problematic as that of a meaning (I shall return to this briefly in §1.4).

It would be odd to claim fallaciousness in the following syllogism as an argument in favour of the first view, rather than as an elaboration of it:

(24) The word 'ja' is a personal pronoun in Polish

(25) The word 'ja' is a sign of affirmation in German

(26) Some word that is a personal pronoun in Polish is a sign of affirmation in German.

On the second view, this syllogism is quite straightforward. Now for Geach, who favours a combination of autonymy and the first view, "a quotation of the Polish word 'ja' will not also be a quotation of the German word 'ja', any more than any other use of the one word is also use of the other."

([43], p.86) But on the second view of word identity, there is no quotation simpliciter - quotation, like any other semantic relation, is language relative. One can still maintain a distinction between autonymy and mere mention on the second view because, even though the same words can recur in many languages, most words do not belong to the English lexicon, and one can only use a word autonomously, ie. to refer-in-English to itself, if it is picked from that lexicon (cp.[69], p.217).

On neither view of word identity is an autonymous view of the mention of foreign words particularly plausible. In writing out his syllogism 24)-26), Geach was writing in English, but since "ja" does not occur in the English

lexicon, he could not have been using it to self-refer-in-English to itself (or as he would prefer to put it, he could not have been using an English word autonomously). If one wants an autonomous theory, one has to say that he either temporarily and subsententially slipped out of English and spoke another language, or he temporarily extended English by taking a foreign word on board. Neither of these two options seems particularly attractive, whichever view you take. The first means that not only would he have created what Kripke calls a 'word salad' of two languages (syntactically), he would, as the second view would put it, be referring-(autonomously)-in-German (or Polish), but predicating-in-English. The second also seems ad hoc, in that the idea of a language being extended by the addition of a foreign word applies only when its usage has become widespread. An account in terms of mimicry is obviously far more plausible when foreign words are mentioned (ignoring the case of conversations between bilinguals, and so on).

Let us now consolidate the position. We have seen that there is self-reference enough in a translational pairing such as 14) to make possible the shift both in the reference to the language and in the reference to words, that would occur in moving to 15). This is not yet to say that it should occur. Secondly, we have the intimation of two distinctions: a semantic one, between mere mention and the autonomous use of an expression, and a pragmatic one, between uninterpreted and translatable occurrences of quotation. This connection between the two seems unproblematic: if an expression is merely mentioned, any use of the whole will be a nontranslatable one, for ex hypothesi we have not

discerned any words to be translated (the question of not preserving reference to the expression, in favour of self-reference, does not arise). So the sentence from the translated language L_1 - German in 14) - should not be translated when passing to 15). We can recognise this to be right on independent grounds: there would be no point in a translation manual if an interpretation of the foreign sentence was already presupposed! But the converse connection is implausible: surely someone could be said to have properly quoted a word, ie. used it autonymously, without that usage having to be translatable. Geach's 'Man' is a noun provides an example, for here no interpretation of the word need be assumed - it is simply being classified syntactically. But this is not the use the translator makes of the sentence from the translating language L_3 ($= L_2$). He is not simply saying: here is a sentence from one language (L_1) and a sentence from another (L_3), and (I conjecture that) they have the same meaning - though what that is is irrelevant. The syntactician, working through his transformations, could say that. But for the translator it is as implausible as it would be for an ordinary English speaker making a direct speech report, eg. with "He said 'It is raining'", to be merely saying in effect, "He produced noises transcribable thus...", without presupposing or conveying what those noises mean. As Burge conjectures, opposition to this rather obvious thought has probably arisen through excessive concentration on rigorous proof-theoretic methods in mathematics, where the mention of an expression is considered in abstraction from its ordinary interpretation ([8], p.149). The translator is using a sentence from his language (in a special way) to convey what one from another language means,

and there seems no good reason for denying that this usage is translatable.

I conclude that as far as the project of interpretation is concerned, manuals of translation have had an undeservedly bad press. Both the truth-theoretic and the translational approaches elucidate meaning by describing what it is to be able to interpret - and in fact we have seen that as far as presentation is concerned, translation manuals have the edge, eg. in the matter of homophony. One might still find constraints on the notion of a theory of meaning to disqualify them from that job. Perhaps one should say that a theory of meaning is not so much one which will impart an understanding of the source language, as that it should state what it is that a person knows in knowing the language. There are then two important questions to be answered: what is it that someone knows when he knows the meanings of the sentences of a language, and is that knowledge amenable to statement by a finite and scrutable theory? For example, if knowing meaning is, paradigmatically, knowing relations between words and world, then the truth-theoretic method is home and dry. Otherwise we may be left wondering, with Dummett, what advantages such a modest theory has over a translation manual⁵.

2: TRUTH-THEORETIC PRELIMINARIES AND THE WALLACE ARGUMENT

We shall consider four fragments of English or the slightly adulterated versions thereof which provide the disciplined, base paraphrases of its ordinary sentences. In our case, each fragment F_i should be encompassed by its successor F_{i+1} , and any truth theory Θ for the latter should build on one for the former, by modifying or generalising it if required. We can imagine the theorist striking out in other directions from the basic fragment F_1 than ours - or even choosing a different basis. In principle, though, all these theories for the different parts of the language should be reconcilable into a consistent, total theory for the whole language (semantic universality permitting).

If the basic fragment contained only names, predicates and the classical connectives, truth, T , could be defined directly as a two-place relation, the first argument place being taken up by ML expressions designating sentences in the fragment and the second by our name for this (sub) language, " F_1 ". But since a theory adequate for English must at some stage deal with demonstratives, I will include in F_1 the simple demonstrative that. As we have already seen, this fundamentally affects the form of the T -sentences. They will take this shape:

(1) For all utterers u , and all times t , $T_c(\varphi, u, t, L) \equiv p$. Truth is irreducibly relativised to the contextual features u and t , since it cannot be conceived to attach to a sentence independently of a particular utterer and time. That is to say, we cannot recover the simple relation $T(\varphi, L)$ by stipulating it equivalent to the LHS of 1) (or anything else).

For simplicity, we shall assume that the reference of a proper name α in F_1 is unaffected by contextual factors. There will be a finite list $\Theta_1 a$ of clauses such as

$$(2) \quad R_c(\text{Jones}, u, t, F_1) = \text{Jones}$$

For the demonstrative, the following would be something like the kind of axiom Davidson has in mind:

$$(3) \quad R_c(\text{that}_i, u, t, F_1) = \text{rvDem}(v, i, u, t)$$

where " $\text{rvDem}(v, i, u, t)$ " abbreviates "the object v which u demonstrates in his i -th utterance of that at t " (cp[101]).

The numerical index is included to cover the possibility of that being used more than once in a single utterance, with different referents. I am ignoring a problem raised by Taylor for 3), that someone might demonstrate more than one object with a single utterance of that. Since this can be considered a case of ambiguity - that should have a single referent - we could perhaps count it as two utterances made simultaneously, and arbitrarily stipulate an ordering of the indices of the two that's. There will be another list $\Theta_1 b$ for the simple predicates, each clause generalised as before for all u and t , though with the assumption that, like the names, the predicates are not imbued with any indexicality. For example:

$$(4) \quad T_c(\alpha \text{ is brave}, u, t, F_1) \equiv R_c(\alpha, u, t, F_1) \text{ is brave.}$$

Recursive clauses for the connectives

$$(5) \quad T_c(\text{Not}\varphi, u, t, F_1) \equiv \text{Not } T_c(\varphi, u, t, F_1),$$

$$(6) \quad T_c(\varphi \text{ and } \psi, u, t, F_1) \equiv T_c(\varphi, u, t, F_1) \text{ and } T_c(\psi, u, t, F_1)$$

complete the definition.

There are basically two ways of extending Θ_1 to cope with the next fragment, F_2 , which incorporates the basic quantifiers for all and for some (cp. eg. [38], section II). I will not consider the various ways of implementing the

Fregean, direct recursion on truth, even though it has the advantage over its rival, the standard Tarskian theory Θ_2 , of harmonising better with the thought that the sentence (or utterance) is the primary semantical unit. Θ_2 introduces the notion of truth relative to a sequence, or satisfaction, T_{sc} . For proper names this extra relativity is idle. We now have, for all sequences s , and all u and t ,

$$(7) \quad R_{sc}(\text{Jones}, s, u, t, F_2) = \text{Jones}.$$

It is the variables which refer relative to sequences, rather as demonstratives assume determinate reference relative to a context; the i -th variable y_i picks out the i -th member of the sequence, s_i :

$$(8) \quad R_{sc}(y_i, s, u, t, F_2) = s_i.$$

Sequence- and context-relative truth T_{sc} is defined for the basic predicates, eg.:

$$(9) \quad T_{sc}(\alpha \text{ is brave}, s, u, t, F_2) \equiv R_{sc}(\alpha, s, u, t, F_2) \text{ is brave}.$$

For existential quantification we have

$$(10) \quad T_{sc}(\text{for some } y_i \varphi, s, u, t, F_2) \equiv \text{for some } y_i T_{sc}(\varphi, s(i/y_i), u, t, F_2),$$

where $s(i/y_i)$ is the sequence which results when y_i evicts the occupant of the i -th place in s and takes its place.

The clauses 5) and 6) will also require relativisation to sequences, to allow for the interpretation of complex predicates, formed with not, and, etc.

The difference between sequence- and context-relativity (between variables and that), is that the former is eliminated. Θ_2 defines truth unrelativised to sequences, T_c , for closed sentences by assimilating them to predicates. As 0-place predicates, they are either true of all objects, or of none; satisfied by all sequences, or none, so

$$(11) \quad (\forall u)(\forall t)(T_c(\varphi, u, t, F_2) \equiv \text{Closed}(\varphi) \ \& \ (\forall s)T_{sc}(\varphi, s, u, t, F_2))$$

In addition to the apparatus for describing the syntax of the OL, the ML must contain enough set theory to describe the requisite properties of sequences, such as

$$(12) (\forall s)(\forall n)(\forall v)(v \in A \rightarrow (\exists s')(s' = s(n/v))).$$

As Davidson says, the aim of deriving all theorems of the form of 1) by itself makes no call for conformity with extensionality or classical logic. "It invites us to use whatever devices we can contrive appropriately to bridge the gap between sentence mentioned and sentence used" ([19], pp. 78-79). But when we come to consider the next fragment F_3 , which contains (sentence modifying) intensional idioms, there is a well-known obstacle to bridging that gap in the shape of what I shall term "the Wallace argument". To accord with previous discussions of the argument, I will omit the indexicality parameters for the moment. By analogy with 5), the natural suggestion for Θ_3 is a homophonic clause for any of these modifiers, schematically represented by " \Box ";

$$(13) T_{sc}(\Box\varphi, s, F_3) \equiv \Box T_{sc}(\varphi, s, F_3).$$

The problem Wallace primarily raised about this concerns the lack of inferential connections between the RHS of 13) and the rest of the theory, adequate for us to continue semantically processing the mentioned OL sentence ([106], p.241). For example, it would be illegitimate to substitute material equivalents established as such by Θ_2 within the intensional context created by the occurrence of \Box in the ML. A way round this would be to adopt a version of the weak rule of inference characteristic of all classical modal systems:

$$(RE) \text{ If } \vdash_{\Theta_3} (\forall v_1) \dots (\forall v_n)(\varphi \equiv \psi), \text{ then } \vdash_{\Theta_3} (\forall v_1) \dots (\forall v_n) (\Box\varphi \equiv \Box\psi).$$

This rule suffices for the substitutivity of provable material equivalents in all contexts. RE) is, of course,

not a rule which any old non-logical theory can help itself to - but its adoption here might be defended on the grounds that Θ_3 is not any old theory. It aims to produce equivalences which are semantical in nature; it would not avail itself of the kind of 'merely factual' equivalences obviously inimical to the use RE). Nevertheless, a more discriminating approach might be thought preferable, one which explicitly strengthened the semantical axioms of the theory, and this is what Wallace proposes: make it deliver strict equivalences, by prefixing its recursive clauses with a \Box . And this leads to the fundamental problem, in making derivable such theorems as

$$(14) \Box(T(\text{Lizzy is playful}, F_3) \equiv \text{Lizzy is playful}).$$

For, supposing \Box to represent logical truth, "it simply is not logically true that Lizzy is playful is true-in- $[F_3]$ English if and only if (materially) Lizzy is playful, unless "true-in- $[F_3]$ English", "Lizzy", "playful", "Lizzy is playful", etc. are all logical constants" ([106], p.223).

This is not the most telling formulation of the point. They don't all have to be logical constants, else not even "If Lizzy is playful then Lizzy is playful" would count. The general problem Wallace is pointing to, the "falsehood objection", does not depend on this particular reading of \Box anyway, it concerns the embedding of truth-theoretic apparatus within the scope of the intensional operator. And this will afflict your truth theory no matter what characterisation of consequence you adopt in the ML, as long as it includes some version of the homophonic axiom 13). Where \Box represents some propositional attitude construction such as Maggie believes that, we would have

$$(15) T_{sc}(\alpha \text{ believes that } \varphi, s, F_3) \equiv R_{sc}(\alpha, s, F_3) \text{ believes}$$

that $T_{sc}(\varphi, s, F_3)$.

Most probably Maggie knows nothing about satisfaction (or even truth, for that matter).

It is worthwhile reflecting briefly on the generality of the problem. 13) is no better when \Box is read as It is possible that. Suppose that in place of φ we have something of the form φ and not φ . The LHS of 13) will be false. But the RHS will be true, given the thought which the Wallace argument is sometimes stated as deploying, that words could have had different meanings. And might have meant what or does. So the whole biconditional 13) will be false. Similar considerations apply to binary sentential connectives such as the subjunctive conditional $\Box \rightarrow$. "If it were both the case that snow is white and not the case that snow is white, the law of non-contradiction would be invalid" is, if intelligible, true; but the correlative sentence of the form $T_{sc}(\varphi \text{ and not } \varphi, s, F_3) \Box \rightarrow T_{sc}(\psi, s, F_3)$ is, by the argument, intelligible and false. The modal thought has a temporal analogue: the expressions of a language change in meaning over time. Reading \Box as a past or future tense operator will therefore also lead to falsehood. (However, it should be borne in mind that relativising the truth predicate to a time - one of the parameters temporarily suppressed - is not readily reconcilable with the use of tense operators in the ML). Finally, we might also mention all those operators which more or less blatantly involve quotation of the embedded sentence, eg. It is written in green ink that.

Since the problem is thus quite general, it might be thought that any solution to it must match it in generality. One policy would be to try to find a way to disown any such

dog's dinner as "Maggie believes that $T_{sc}(\text{Schmidt ist böse}_2, \text{s, German})$ " that might be served up. For although this could occur as part of a consequence of a truth theory for German, it would not be a testable consequence, the argument could run. The only testable consequences of such a theory for a natural language are the target T-sentences, and these either mention an OL sentence in its entirety, on the LHS, or else fully translate it, on the RHS. All the rest, it may be said, is just instrumental machinery for the generation of these biconditionals.

Something faintly reminiscent of these ideas can be found in the writings of Davidson and McDowell (eg. in [24] and [68]). They would agree that the evidence for an interpretative truth theory does not come at the subsentential level; although in the postulates such as θ_1a and θ_1b relations of reference and satisfaction are asserted to hold between words and world, neither the axioms nor anything intermediate between them and the final T-sentences are confirmed directly. They are tested through the ability of the T-sentences they generate to serve the general project of making sense of the speakers of the OL. But there is hardly a kind of instrumentalism here to come to the defense of 13). The concern of Davidson and McDowell is with the direction of confirmation in a truth theory in the context of the radical interpretation of some tribe, from theorems to postulates, not with the truth status of the latter. It is central to their conception of such a theory that its axioms be intelligible and true statements - eg. of an "eminently learnable and forgettable relation between an English word and a set of men" ([40], p.xi). It is simply that they do not receive independent confirmation.

Without some strong and implausible instrumentalist principle to effect a split between the directly and indirectly confirmed sentences, in regard to their content-bearing status, one would have to give up the idea that even the finished theorems state anything, if pursuing this line of approach. And if one is giving up the ambition of stating the truth conditions of OL sentences, one might as well have adopted the translational format from the outset. For compare 1.1.11, which is immune to the Wallace argument. For the genuine truth theorist, there appear to be only two options. Either he finds an acceptable way of rendering the construction within an essentially extensional ML, or he finds a way of making the (apparent) intensionality acceptable in the ML (perhaps leading to a more profound understanding of extensionality anyway - cp. [80], pp.474-5).

3: THE PARATACTIC THEORY

It is natural to think that one way a Davidsonian theory of meaning would accomodate modal locutions would be by suitable adaptation of Davidson's own paratactic account of the logical form of indirect discourse sentences. (The possibility has been suggested in the literature though not, to my knowledge, explicitly by Davidson himself.) According to the O.E.D., parataxis is "the placing of propositions or clauses one after another, without indicating by connecting words the relation (of co-ordination or subordination) between them". By such a definition, even an operator treatment of a sentence such as

(1) Galileo said that the earth moves

in terms of the application of the operator Galileo said that to the content sentence

(2) The earth moves

will count as what we might call "the minimal paratactic theory". But what has become known to philosophers as the paratactic theory - Davidson's account of 1), offered as a prototype for the "whole unholy array of attitude-attributing locutions"¹ - incorporates a more substantial body of doctrine concerning their form and content. I shall concentrate on the prototype, with just a few words on a paratactic theory for modality, for the difficulties which seem to me to arise concern more the basic proposal than its extension to other cases. It will be useful to list the most distinctive components of Davidson's theory as clearly as possible in that, should they prove not to be jointly satisfiable, we may ask whether a modification of the theory can be made which does not lapse into the minimal theory.

a) A proposal about the logical form of propositional attitude sentences, like any other logical form proposal, must be capable of incorporation into a semantical theory, such as Tarskian truth theory adapted for natural language ([15], p. 144).

b) Both the components discerned as paratactically combined are complete sentences. Correspondingly, an utterance of a propositional attitude sentence, such as 1), really consists of two distinct utterances, of the two sentences. In addition to the content sentence we have what for want of better terminology I shall term the "mood setter"², ie. for 1) it is

(3) Galileo said that.

3) is a sentence made up from a two-place predicate and two singular terms. Mood setter and content sentence are punctuated by an ordinary full-stop.

c) The singular term that in the mood setter is a demonstrative. The true demonstrative nature of the that in a construction such as 1) is hidden to us both by the proximity of its reference (this would be better), and by its anaphoric nature - unlike the more familiar that, it requires no supplementation with an act of demonstration. In any context in which 1) is uttered, the reference of its that is fixed. Points b) and c) are corroborated by the use of an expression such as Whistler said that as a complete sentence, with the that referring demonstratively to something, eg. a previous utterance ([15], p. 143).

d) The notion which the predicate said expresses is the ordinary notion of saying the same which we employ when reporting the words of another. The inference from 1) to (4) Galileo said something

has been shown to be valid - basically a case of existential generalisation. After all, 1) is true iff Galileo said something - not necessarily 2), for our concern is with indirect quotation, and there may be changes in tense, word order and so on. More importantly, Galileo didn't have to speak my language in order for me to utter 1) truly. What is required is that the utterance he made have the same content - according to our ordinary, shifting standards of samesaying - as an ordinary English utterance of the content sentence 2). Let us introduce said' for this oratio obliqua notion, to be distinguished from the said'' of oratio recta, where the particular words spoken are at issue. Thus we can say:

- (5) Galileo said' that the earth moves iff $(\exists v)(\text{Galileo said''}(v) \text{ and } v \text{ matches in purport the following: The earth moves})$.

This is by way of elucidation, not analysis (cp.[22],p.39). The notion of synonymy which samesaying involves is not to be analysed by the semantic theory. (The main reason for this stems from the distinction between matters of (logical) form and matters of analysis - recall constraint C) from §1.1, and cp.[13],pp.316-17)..

e) The reference of that is an utterance - not a sentence or a proposition, for example. Consequently

f) The connection between the two components of a propositional attitude sentence is not logical or semantical, but pragmatic³. Or rather their utterances are so related; the utterance of 3) containing a reference, on any occasion of utterance of the whole, 1), to the utterance of the content sentence, 2) ([15],pp.150-51).

g) All the words function in the way they normally do.'

In particular, the words in the content sentence do not refer to their senses ([15], pp. 151-52), i.e. the theory vindicates Davidson's "objection to meanings in the theory of meaning ... not that they are abstract or that their identity conditions are obscure, but that they have no demonstrated use"⁴. With a), this means that

h) All the words in a propositional attitude sentence will get the treatment they standardly receive from the truth theorist. Standard treatment by the formal theory is the criterion of normal function (g). The theorist recovers his pre-Fregean innocence (cp. §2.1).

i) When 1) is asserted, only the mood-setter is asserted. There is no need for anyone to concur with Galileo in reporting him - though if one added and I agree to 1), and asserted the whole, one would effect an assertion of the content sentence. That the content sentence is uttered unasserted is as unsurprising as its unasserted utterance when eg. disjoined with another sentence. If 1) is to be used in an argument, the term the earth does not, strictly speaking, occur in it as part of a premise, but only in something referred to by the premise, which is 3) ([26], p. 10). It is not available for substitution on the basis of a true identity - substituting for the earth will change the referent of the that ([15], p. 151).

The extension of the account to modality is fairly straightforward. For something of the form It is necessary that ϕ it is obvious. When we have an utterance of something like

(6) The earth cannot move

understood as expressing a conclusion about the dictum that the earth moves - that it is impossible - we can construe it

as the joint and simultaneous utterance of 2) and

(7) That cannot be.⁵

Suppose that at t_1 Jones utters I am brave. Later, at t_2 , Smith reports this with an utterance which we paratactically punctuate as follows:

(8) Jones said' that. He was brave.

The tense of the content sentence conveys the temporal content of Jones's utterance well enough for ordinary purposes (supposing enough of a gap between t_1 and t_2), for Jones said that he is brave at t_1 , and $t_1 < t_2$. He in Smith's mouth at t_2 refers to Jones, just as I does, in Jones's mouth, at t_1 . Smith's use of 8) would normally have been acceptable if Jones had used Jones instead of I.

Whether it would still be acceptable if Jones had used a descriptive phrase uniquely true of himself - ie. whether ordinary speakers would count that as the same thought as expressed by the content sentence of 8) - would depend in part on Jones's purpose in clothing his thought in that way, and on Smith's in baring it. This suggests that when we come to consider quantifying in along the lines Quine suggested

(9) Something is such that Galileo said' of it that. It moves

our ordinary standards of indirect reportage will tolerate the shift to It.

How does the paratactic proposal work as part of the truth theory for English? Let us start from the initially plausible assumption, consistent I think with what has been said so far, that, as one would expect of any theory about the semantical composition of intensional sentences, it will eventually lead to their interpretation as complex wholes.

That is, we shall expect to find some condition to match
 (10) $T_{sc}(\text{Galileo said' that. The earth moves, } u, t, F_3)$,
 one which will enable us to say what it was that u said'
 that Galileo said'. The main question then is: what treat-
 ment should be provided for the demonstrative that? It is
 natural to assume that the standard clause 1.2.3 would be
 proffered; indeed Davidson hints at this in "Truth and
 Meaning" (p.320). The indexing of occurrences of the demon-
 strative is obviously essential if the theory is to accom-
 modate the embedding of one propositional attitude construc-
 tion within another. Ignoring the possibility that Galileo
 said what he did at midnight, in the bathroom,... and the
 problems for logical form which attend this, we might expect
 to find as the ML condition matching 10):

(11) Galileo said' $\text{rvDem}(v, 1, u, t)$. The earth moves
 Clearly this is no good. No connection between the mood
 setter and content sentence has been effected in the ML.
 And so, for example, if the theorist accepted that an
 utterance of 1) was true, relative to some u and t , he would
 be forced, by detachment from the 10)-11) biconditional, to
 assert both that Galileo said the first thing demonstrated
 by u at t , and that the earth moves. This is simply a
 consequence of the fact that the theorist must be prepared
 to assert his biconditionals, and that there is no demon-
 strative reference in the ML to "The earth moves" and
 consequent attribution of the saying' of it to Galileo. 11)
 is inadequate whether we arrive at it through attempting a
 semantical treatment of the punctuation mark

(12) $T_{sc}(\varphi \cdot \psi, u, t, F_3) \equiv T_{sc}(\varphi, u, t, F_3) \cdot T_{sc}(\psi, u, t, F_3)$
 or whether we derive the truth conditions of 2) and 3)
 separately and then fuse them as deep structure constituents

of a single surface form.

It may seem that I have laboured a rather obvious point - but let us consider the nature of the inference from 1) to 4). According to the paratactic theory, it has the schematic form

$$\varphi(a,b).\psi \vdash \exists v\varphi(a,v)$$

Obviously the first step in this inference is not existential generalisation, for then the problem of the assertion of the content sentence would arise again. For example, from 8) we would get:

(13) Jones said' something. He was brave

Although Jones spoke a language, his bravery has been a matter of some controversy amongst scholars. Rather we must assume that there is a rule which allows us to drop ψ , the content sentence, ie. that we can go from 1) to 3). If that was treated as a constant, anaphorically referring forward to whatever follows it, the inference would be invalid. We have to see the that as demonstratively referring back to the previous utterance of the content sentence, rather as in Wilde's famous remark. The only alternative is to see the inference as proceeding directly from 1) to 4), but this hardly favours Davidson's paratactic punctuation. It supports the rival theory which sees the whole that-clause as a name, the something which Galileo said, an intensional entity.

Suppose we forget about the inference from 1) to 4) (and the explicitly demonstrative cases such as Wilde's remark) - what are the prospects for an alternative treatment of that? As already noted, the central occurrences of that which are succeeded by a content sentence have a certain contextual rigidity, so we could aim to treat them as

constants. According to Davidson, "No logical errors result if we simply treat demonstratives as constants; neither do any problems arise for giving a semantic truth definition" ([13], p.319). Distinguishing this that' and the other, the axiom which suggests itself is:

$$(14) R_{sc}(\text{that}', u, t, F_3) = \text{that}'.$$

After all, if it is a demonstrative which holds together complete propositional attitude constructions, it is only natural to expect that if they are to be interpreted as wholes, it will be used in the ML in stating their truth conditions. In his comment on [15], Quine seems to take the paratactic theory this way. But what is stated in this axiom 14)? Presumably that' must mean something which can be elucidated in such terms as these: the utterance which follows immediately. If this is so, then the statement of 14) will be false, in general - depending on what the theorist follows it with. At best, it will be truth-valueless, if he manages to observe a suitable period of silence. To put it another way, if the theorist is really using that' in the ML, rather than mentioning (using autonomously) it, the object he demonstrates enters into the truth conditions of every statement involving that' (cp. 81.1).

In the absence of any suitable homophonic axiom for the that-clause that', it would seem that the paratactic theorist is forced back to treating said' that as a single unit - as a sort of intransitive verb, if the paratactic punctuation mark is retained. A distinctive feature of Davidson's version is thereby removed: the claim to have discerned a singular term following said'. And now we are back to 1.2.13 in essence, and the Wallace argument is reinstated. There is a general dilemma facing the truth

theoretic incorporation of parataxis: either the content sentence is not suitably linked to its predecessor in the ML, or it is, and the theory has the means to link sentences in a manner found problematic by the Wallace argument⁶. Assuming there to be a rule permitting substitution of proved equivalents between full stops, we could derive on the RHS: (15) Galileo said' that. $T_{sc}(\text{The earth moves}, u, t, F_3)$.

It would seem, therefore, that the paratactic theorist has no alternative but to accept the first horn of our dilemma - that we should drop the assumption we have been proceeding under that it makes sense to talk about the meaning or truth conditions of such sentences as 1), in toto. He may query the assumption that the period is in some sense a semantic (or pragmatic - cp. fn.3) constituent. After all, the ordinary full stop is simply a device of punctuation. (Moreover, the loss of an axiom like 12) - which anyhow would have to be accompanied by a characterisation of this connective in the ML - would only foil the accomodation of quantifying in etc., as in 9), if we have to construe such cases of anaphoric reference accross a full stop along the lines of quantifier and variable. That is by no means certain.) The paratactic theorist may reason that there is no way of directly stating the truth conditions of fragments of discourse longer than one sentence, and that 1) is just such a fragment. All one can do is state the truth conditions for individual sentences; for example (16) $T_{sc}(\text{Galileo said' that}, u, t, F_3) \equiv \text{Galileo said'}$
 $\text{ivDem}(v, 1, u, t)$.

and similarly for the content sentence. (Alternatively, he could link 10), on the LHS, with the RHS of 16), which by his lights would be true, but by no one's lights would be

interpretative.) The intuitive unpalatability of a view which denies that we can ascribe truth or meaning to whole propositional attitude sentences should have been allayed by the informal statement of the theory in a)-i) above. From the informal statement of the theory, one might think it implicitly a quotational theory, since the reference to the content sentence (the sentence being one component of the utterance - cp. [69], section V) is exploited in accounting for failures of substitutivity. From the formal version of the theory, one can neither say that it is nor that it isn't a quotational treatment of propositional attitude sentences - it is a dissolution of them. There is no way one can recover an interpretation of the complete sentence, on this theory. Perhaps afficionados of the position regard all this as a statement of the obvious, yet I think it worth emphasising if only because it requires a major revision (or U-turn, in popular parlance) in Davidsonian policy. It means reneging on the promise to state the truth conditions of sentences as they occur in the OL. (This is not just a quibble about one small point being omitted from the OL - as the variety of intensional constructions to be considered in chapter 2 should make apparent, it will involve quite a carve-up.)

Before passing final judgement on the paratactic theory, we should consider a theory which McDowell has promoted as an explicitly quotational variant of it ([69]). Davidson comes to his paratactic theory by way of a rejection of a quotational theory of indirect discourse due to Quine ([15], pp. 148-49). McDowell rescues the quotational account from the two distinct but overlapping objections which Davidson raises, only to attack it with some objections

of his own. Davidson's first objection: any theory which construes the indirect reporter of a speech as not having uttered the sentence he is reporting fails to make him a samesayer with his subject, and thereby fails. McDowell replies that as long as he specifies the sentence somehow (eg. by quoting it), that will suffice. Samesaying holds between potential utterances ([69], section V). Davidson's second: to quote a sentence is not to utter it. Given a spelling account of quotation, the objection is fair enough; but as we have seen, that is not the only account, nor the best. By backing the autonymous theory of quotation, McDowell dispatches the second objection. A line of thought he then pursues can be put in the form of a question: since we can utter quotations, and utter them unasserted, and since the content sentence of an oratio obliqua sentence, rendered paratactically à la Davidson, is uttered unasserted, what is to stop us thinking of it as quotational? If quotations are uttered, we can use said' in a quotational analysis of the indirect report 1):

(17) Galileo said' 'The earth moves'.

And now, contrary to McDowell's final intentions, if we assume the autonymous theory - which enables us to dispose with the quotation marks - and combine it with paratactic representation of logical form, this converts into the familiar shape of

(18) Galileo said' that. The earth moves.

McDowell's answer to the question just posed in his name is that if there is such a thing as a successful paratactic quotational theory, it cannot differ in essence from Davidson's original theory. His reason for maintaining this stems from consideration of the behaviour of indexical

expressions within quotes. To return to a previous example, suppose that Smith had used, in place of 8):

(19) Jones said' 'I am brave'.

I here refers to the word I, though 19) conveys the information that a particular token of I was used by Jones to refer to himself. If instead Smith says:

(20) Jones said' that I am brave

the I refers directly to Smith. If an indexical is embedded within the quotation marks of an oratio recta sentence such as 19), its reference no longer relates back to the context of utterance of the whole sentence, rather it relates indirectly to the context introduced as that in which the quoted material is or was produced. This "sealing off" effect of quotation marks might appear at first sight to vitiate any quotational construal of an indirect discourse construction containing indexical words, such as 20). But of course it would not be 19) that is proffered as an analysis of 20), but something involving said':

(21) Jones said' that. I am brave

When, on the paratactic account, Smith utters 21), he describes Jones as saying something matching the content of Smith's utterance of the second sentence, the quotational I am brave. This is very different to 19). Since the context of this utterance has Smith as its speaker, we can if nothing else stipulate that this (notion of saying') requires of Jones that he said something matching Smith is brave, ie. I am brave as it would be uttered by Smith, the result desired for (Smith's utterance of) 20). At this point it is natural to want to press the query: but in what sense is this really a quotational analysis - a sense which makes it a contender to Davidson's own proposal (cp.[69],

pp.228,229). The answer is that it must show up in the axiom for said'. In the ML the theorist will need to employ a predicate "said*" which will relate the quotation of the content sentence he follows it with on the RHS to the use which his subject u would make of it:

$$(22) T_{sc}(\alpha \text{ said' that } \varphi, u, t, F_3) \equiv R_{sc}(\alpha, u, t, F_3) \\ \text{said* that}(\varphi, u, t).$$

Smith's utterance of 20) at t , for example, should be true iff Jones said* that(I am brave, Smith, t), ie. iff Jones said something matching the content of a normal utterance of I am brave by Smith at t , had he made it (rather than the autonymous use which, according to the theory, he then made of that sentence.)

22) is a quotational treatment, in involving reference to a sentence. It also amounts to an implementation of Davidson's idea of reference to an utterance in that, in order to cope with the occurrence of indexicals within the content sentence, we have to keep track of the other two components of an utterance of it, the utterer and time. In designing a theory for English in English, we have employed the identity mapping for the content sentence φ , but as Church would be happy to remind us, 22) would not be an interpretative style of axiom for a foreign OL. However, it should be evident from §1.1 that I would not consider this a decisive objection (Burge develops the case for translatable quotation in [8], which I drew on in that section, with special reference to quotational treatments of oblique contexts.) We might therefore get round the problem by supplementing 22) with a rule which allows that when $TR(\varphi, L) = \psi$ (L is the OL and ψ is a complete translation of φ from it to the ML, ie. contains no occurrences of TR), we can substitute

ψ for φ^7 . So if in interpreting Schmidt the counterpart of 22) for German served up on the RHS: Jones said* that(Ich bin tapfer, Schmidt, t), we could substitute "I am brave" for "Ich bin tapfer".

If a quotation is a structured singular term made up by words and concatenation, the full-stop in 18) and 21) would, ungrammatically, be connecting a sentence and a singular term. But McDowell has a complaint against the treatment of quotations as terms. It is ([69], p.222) that in so doing the theory will not be able to handle such sentences as:

(23) Geach says that they 'logically attach to the subject'. Obviously, if α is a singular term and they occurs non-autonomously, they α will be ungrammatical. But what this shows is that appearing as quoted material cannot be the main function of the words between the inverted commas. It is clear that these words are being used twice over in 23), a feature reminiscent of the Giorgione example (1.1.20): once, in the way words are normally used in the content sentence of an oratio obliqua construction, and secondly, as tokens of the very same words that Geach is alleged to have used. The appearance of grammatical impropriety only arises if one fails to distinguish the oblique and the direct aspects of the sentence. 23) might be represented as:

(24) Geach says that. They 'logically attach to the subject' where someone says (unprimed) a demonstrated sentence iff he says' it and he does so in part or in whole by using any of the indicated words which occur between inverted commas, if such there be. We thus let saying coincide with saying' when the content sentence contains no inverted commas. The inverted commas are merely special markers for the oratio

recta component of saying - they are not quotation marks in the sense that Geach would be described, in 23), as uttering a sentence itself containing quoted material. On the paratactic quotation account, the real quotation marks occur, invisibly, enclosing the whole of the content sentence.

In conclusion, I think we can say that Davidson's informal paratactic theory leads to at least three different proposals. Firstly, consideration of indirect discourse serves to remind us that the object of the propositional attitude must match, by some standard of accuracy, the gist of the content sentence used by the reporter of the attitude. This suggests that we may distinguish a liberal notion of saying, believing, etc. for which the likes of 15) are acceptable. By the lights of the theorist, who after all is the one who would utter 15), Galileo did say something tantamount to: The earth moves is true relative to u , t and F_3 , since that, on a good theory, is tantamount to this: The earth moves. But if that is one kind of propositional attitude, there must be other stricter notions which cannot be treated in that manner. For these the paratactic theory offers us a choice: either split them in two (never the twain to meet), or treat them as disguised quotational contexts.

4: OPERATORS IN THE ML

The kind of theory to be studied in this section is typified by an appeal to special features of certain modifiers, such as the alethic modalities, as exempting them from the general unacceptability of 1.2.13. After all, there should be some notion of 'semantic necessity' for which necessitation of the axioms and theorems of a semantic theory is an acceptable rule, as much as there is eg. physical necessity, as necessity relative to the laws of physics. Whether this notion of semantic necessity should be identified with analyticity is another matter, though (think of the general case where ML and OL are two different natural languages). But at the very least, we should be able to give a theory of truth for the special case where this theorist's notion of necessity is imported into the OL¹.

If special pleading for the homophonic treatment of certain traditionally intensional operators is to be persuasive, it must ultimately be accompanied by an alternative strategy for those more blatantly intensional locutions for which 1.2.13 is conceded to be unsatisfactory. It would still be preferable to treat the exceptions according to 1.2.13 rather than the more general alternative, for homophony is to be preferred wherever possible. The idea that metaphysical necessity might also qualify as such an exceptional case should not surprise us too much, for we have become familiar with the idea of its exceptional status with regard to the substitutivity of coreferential proper names within its scope. The difficulty alleged, which concerned the possible falsity of Hesperus is Hesperus, for instance,

hinged not on the possibility of the nonlinguistic facts changing for the worse, but of the meanings involved so doing. A natural defense of the use of metaphysical necessity to interpret itself, using 1.2.13, would therefore consist in a justification for ruling this hypothesised fluctuation in meaning illegitimate.

An example of this response is provided by Gupta's appeal to two different notions of truth ([48], especially p.453). There is room enough for this manoeuvre, because the truth theory to which 1.2.13 is added as an axiom will only have served to define the extension of the concept of truth, not its intension. The first truth predicate, T_1 , is such that $\Box T_1(\varphi, L)$ holds just in case for every possible world w the proposition p which φ expresses in w is true in w . For T_2 , $\Box T_2(\varphi, L)$ holds iff the p which φ actually expresses is true in all worlds w . The basic idea is that the required constancy of meaning can be attained by giving a theory of T_2 truth. Gupta does not supply us with a reading of T_1 and T_2 which eliminates mention of worlds, so it might be wondered whether this is a genuinely homophonic approach. We have to understand the necessity operator in the NL as binding the world variables in a third truth predicate, $T_3(\varphi, w_1, w_2, L)$, where w_1 is for facts, and w_2 is for meanings. $\Box T_1(\varphi, L)$ is $(\forall w)T_3(\varphi, w, w, L)$; and $\Box T_2(\varphi, L)$ is $(\forall w)T_3(\varphi, w, @, L)$. But the idea is clear enough: if the truth theorist is to utilise 1.2.13 as his axiom for necessity, he should understand the truth predicate it involves in the following, restricted sense: truth with respect to actual meaning. He can of course be prepared to construe the predicate thus in advance of knowing what these particular actual meanings should turn out to be. The same idea could

be used for temporal modifiers: we can concentrate on current English, for we are to construe the truth predicate as truth with respect to present meanings.

In the presence of a challenge that ϵ_3 can't work, because the property of sentences φ it requires to be instantiated - of its being a necessary fact that φ is true - can never be exemplified, Gupta's answer is, crudely, that ϵ_3 can work because we can understand truth in such a way that the property does have the required instances. Some have sought to provide a deeper motivation for the acceptance of such theories as ϵ_3 , by arguing for the necessity of the likes of $T(\text{Hesperus is Hesperus}, F_3)$ on conceptual grounds. The argument is that those words could not mean anything different in (F_3) English, for the very identity of that language (or fragment) depends on the meanings of the words which belong to it (cp.[5],p.84,[80],pp.477-78). The defense would also apply to a species of necessity weaker than, but not exclusive of, metaphysical necessity, such as "it is at least causally necessary that" (cp.[81],fn.18). Since what we have here is an essentialist claim about the identities of languages, it will be convenient to pause to introduce the final fragment F_4 , containing the lambda notation often thought suited to the expression of such claims, and make a detour through some truth theories for it.

Lambda abstracts are usually understood to be singular terms, denoting either sets or properties. So whereas man(Caesar) expresses the thought that Caesar falls under the concept man, in $(\lambda v_1)\text{man}(v_1)$ [Caesar] the square brackets express a different metaphysical relation, that of having or participating in, holding between the two entities denoted by Caesar and $(\lambda v_1)\text{man}(v_1)$. We can envisage a

theory Θ_4 which accomodates λ -abstracts by employing the new mode of semantic combination in the ML:

$$(1) \quad T_{sc}((\lambda v_i)\varphi[\alpha], s, F_4) \equiv (\lambda v_i)T_{sc}(\varphi, s(i/v_i), F_4)[R_{sc}(\alpha, s, F_4)].$$

Multiple abstracts can be defined in the usual way:

$$(2) \quad (\lambda v_1) \dots (\lambda v_n)\varphi[\alpha_1, \dots, \alpha_n] =_{df} (\lambda v_1)(\dots((\lambda v_n)\varphi[\alpha_n]) \dots)[\alpha_1].$$

So these can be interpreted by repeated applications of 1).

If, as in Θ_3 , we let \Box attach to both open and closed sentences, the logical form of

$$(3) \quad \text{Caesar necessarily is a man}$$

understood as an attribution de re, can be revealed as either

$$(4) \quad (\lambda v_2)\Box(\exists v_1)(\text{man}(v_1) \ \& \ v_1 = v_2)[\text{Caesar}], \text{ or}$$

$$(5) \quad (\lambda v_1)\Box\text{man}(v_1)[\text{Caesar}],$$

depending on how much deference you pay to the words is a in 3) (see below). Choosing the latter for simplicity, we know by 1) that for any sequence s 5) is true-in- F_4 , relative to s , iff

$$(6) \quad (\lambda v_1)T_{sc}(\Box\text{man}(v_1), s(1/v_1), F_4)[R_{sc}(\text{Caesar}, s, F_4)].$$

We also know, by 1.2.13, that

$$(7) \quad T_{sc}(\Box\text{man}(v_1), s(1/v_1), F_4) \equiv \Box T_{sc}(\text{man}(v_1), s(1/v_1), F_4).$$

By application of Θ_4 b (the list for predicates), and 1.2.8, we know that the following are equivalent (for all v_1):

$$(8) \quad T_{sc}(\text{man}(v_1), s(1/v_1), F_4); \text{man}(R_{sc}(v_1, s(1/v_1), F_4)); \text{man}(v_1)$$

Using RE) and predicate logic, we may infer

$$(9) \quad T_{sc}(\Box\text{man}(v_1), s(1/v_1), F_4) \equiv \Box\text{man}(v_1).$$

In order to utilise this equivalence, we need a rule akin to one proposed by Grandy:

$$(RA) \quad \text{If } \vdash_{\Theta_4} (\forall v_1) \dots (\forall v_n)(\varphi \equiv \psi), \text{ and if } \neg (\lambda v_1) \dots (\lambda v_n)\varphi \text{ then infer: } \neg (\lambda v_1) \dots (\lambda v_n)\psi.$$

We can then substitute RES for IHS, of 9), in 6), which with

the axiom for Caesar, gives the correct truth conditions of 5). Note that it is the presence of such a rule as $R\lambda$, in a theory which does not take as axiomatic any contingent, non-semantic facts - in contrast to a theory which adopts eg. a rule similar to $R\lambda$ but without the condition that the material equivalences should be provable within the theory - which principally signals that the denotations of the λ -abstracts are being treated as intensional entities.

If the de dicto/de re distinction is to be maintained, the classical principle of abstraction, $(\lambda v_i)\varphi[\alpha] \equiv \varphi(v_i/\alpha)$ (where (β/α) is the result of uniformly replacing all free occurrences of β in φ by α), must be replaced by some more restrictive principle (cp. [96]). We can then distinguish between

(10) $\Box T(\text{Hesperus is Hesperus}, \text{English})$,

(11) $(\lambda\varphi)\Box T(\varphi, \text{English})[\text{Hesperus is Hesperus}]$, and

(12) $(\lambda v)\Box T(\text{Hesperus is Hesperus}, v)[\text{English}]$.

Prima facie, things stand as follows. 10) is the kind of thing which might crop up on the RHS of some consequence of Θ_4 . 11) expresses a claim concerning the sentence Hesperus is Hesperus, that necessarily it is true in English. If this is an essentialist claim, it is a rather dubious one. An essential property is traditionally understood as one on which the very existence of its possessors depends; eg. manhood's being essential to Caesar, as the impossibility of any attempt to conceptually sever the property of being human from Caesar on pain of obliterating him. A property's being essential to an individual can be elucidated, in possible worlds vocabulary, in terms of the individual's having the property in any world in which it exists. This means that, according to 11), in any world in which Hesperus is

Hesperus exists, there English must exist for that sentence to express a truth in. On the view of word identity granted the most plausibility in section 1.1, this is false: the existence of an expression transcends that of any particular language it happens to occur in. The sentence could have existed as a meaningful string in another language, even if English had not come into existence. (A fortiori its having the True as its extension in English is not essential to it.) Finally, 12) would appear to represent the kind of conceptual truth involved in the proposed deeper defence of the use of 1.2.13 - even though the inferential route from it back to 10) cannot be considered to be generally valid (just because it moves the singular term within the scope of the modal operator).

Before attempting to evaluate this situation, it will help to clarify the nature of essentialism if we contrast Θ_4 with a rival theory, Θ_5 , promoted by David Wiggins (see especially [109]). Wiggins does not hold that a unitary account of de dicto and de re necessity is guaranteed from the outset; not wishing to prejudge the issue in favour of unification by adopting one expression for both, he adopts \Box for the former and Nec for the latter. The purpose of Nec is to modify λ -abstracts, which it could do in two ways. If λ -abstracts are terms, Nec can be taken either i) as a 1-place operation constant, forming new terms from old, or ii) as an expression which turns them into predicates. Representative of i) is:

(13) $\text{Nec}(\lambda v_1)\text{man}(v_1)[\text{Caesar}]$

and of ii) is $\text{Nec}(\lambda v_1)\text{man}(v_1)(\text{Caesar})$. A further option iii) would be the introduction of another abstraction operator, $\text{Nec}(\lambda v_1)$, not as a rival to lambda, but as an intensional

cousin of it. Option i) obviously allows Nec to be iterated directly, ie. NecNec will be a permissible combination, and it is this (in Peacocke's generalised theory for Wiggins's Nec, [79], p.314) which signals most clearly that this is the option favoured.

According to the title of one of his papers on the topic, Wiggins is more concerned with the logical form of essentialist's claims - with the correct formulation of a philosophical thesis, than with the analysis of occurrences of de re modality in ordinary English. Nevertheless, Wiggins appeals to the existence of ordinary de re locutions - rather than to possible worlds - to justify the intelligibility of essentialist claims and, crucially, the bona fide nature of the modified predicate $(\lambda v_1)(\text{Nec}(\lambda v_2)(\lambda v_3)(v_2=v_3))$ [Hesperus, v₁] in his de re version of the Barcan derivation of the necessity of identity of Hesperus with Phosphorus (cp. [110], Longer Note 4.11). To properly demonstrate this appealing move, it should be shown how the traditionally metaphysical distinction between de dicto and de re surfaces in the vernacular. Traditionally, if a proposition is deemed necessary, this is de dicto necessity. It relates to the mode of truth. De re modality relates to the mode in which an object possesses a property: necessarily or essentially, actually, possibly,..... But there is an alternative construal. Compare the following, as candidate elucidations of 3):

(14) Caesar necessarily-has the property of being a man,

(15) Caesar has the property of being necessarily a man.

It may not be clear at first if there is any distinction here - what, eg., should we make of the property ascribed in 15) if not the necessary-having of the property of manhood,

à la 14)? Wiggins writes of "treating necessarily as a modifier of predicates or the copula" ([110], p. 107), and nowhere appears to credit the existence of a distinction in this, a distinction which could be described as that between property and copula modification. Sometimes he takes the line I called the traditional one, describing the role of Nec in 13) (his version of 3)) thus: "necessarily as in effect qualifying the copula and signifying the mode of inherence ... 13) says that Caesar necessarily has

$(\lambda v_1) \text{man}(v_1)$ " ([112], p. 37). 14) accords with this account, ascribing the property of necessity to the relation of possession or instantiation, as holding between the individual Caesar and the property of manhood. By contrast, in 15) we have a modal property categorically predicated of an individual. Compare now the elucidation of necessary identity, " $\text{Nec}(\lambda v_1)(\lambda v_2)(v_1 = v_2)$ ", or that relation which any r and any s have iff they are necessarily identical."²

Is the distinction between property and copula modification metaphysically excessive enough for a homophonic theory of ordinary English to pass over without acknowledgment? Consider the following three grades of modal involvement (not Quine's) in this sentence:

(16a) It is possible that the man with the hat is a candidate for the presidency.

b) The man with the hat is possibly a candidate for the presidency.

c) The man with the hat is a possible candidate for the presidency.

(The three grades are also available for an action sentence such as: The president has stopped a nuclear war.) Each of these expresses something different. Thus 16c) says,

concerning the man with the hat, that he is something: a possible candidate for the presidency, whereas 16b), though still about that man, does not assert that he definitely is something. The distinction is not simply adjective against adverb: to say that x has destroyed a possibly fatal tumour is not the same as to say that x has possibly destroyed a fatal tumour. Moreover, since both of these occurrences of possibly are, intuitively, epistemic, it is not clear that a charge of equivocation over the kind of modality involved would have much relevance here.

I do not pretend that with one operator Poss the differences between 16b) and 16c) could not be properly expressed, provided that proper heed is taken of its interaction with the words is a. For example:

(16b') $\text{Poss}(\lambda v_1)((\exists v_2)(\text{CP}(v_2) \ \& \ v_2 = v_1))[\text{the MH}]$.

(16c') $(\exists v_2)(\text{the MH} = v_2 \ \& \ \text{Poss}(\lambda v_1)\text{CP}(v_1)[v_2])$.

But this only serves to emphasise that, construed under option i), these predicate modifiers basically fulfil the function of property modification. Copula modification is achieved derivatively, by forming as a predicate in 16b') be a candidate for the presidency. This means that one can no longer afford to disregard the difference in surface structure between Caesar is necessarily a man and Caesar is a large man. To construe 3) according to 13) is, as we might say, to make the necessity fall within the scope of the having. This becomes apparent if we reformulate 13) in a more traditional manner: $\text{Have}(\text{Caesar}, \text{Nec}(\lambda v)\text{man}(v))$ (where "Have" makes explicit the metaphysical relation expressed above by the square brackets.) This is all right for the adjective large, but the idea that appeal to natural language promotes the use of Nec hardly makes it relevant to

the issue of essentialism, since the English word which occupies this position, necessary, expresses a low grade species of necessity - something like "required for" (cp. "Have you got the necessary equipment?"). Moreover, iteration of Nec produces a term, NecNec(λv)man(v), which with some charity and imagination can be interpreted, but is hardly something an essentialist would have a use for³.

If we leave on one side property modification for the moment, the possibility arises of a new and potentially simpler representation of essentialist claims. Instead of the modality attaching to arbitrarily long abstracts $(\lambda v_1) \dots (\lambda v_n) \varphi$ to form modified relations, the mode of each individual's participation in a relation must be assessed and expressed separately. Using NEC for this form, we might have something like:

(17) NEC(λv_1)NEC(λv_2)($v_1 = v_2$)[Hesperus, Phosphorus].

It will be helpful to make a comparison at this stage. We must think twice about allowing willingly to modify a relation, as in this version of Lakoff's example

(18) Willingly(λv_1)(λv_2)(v_1 sacrifices v_2)[the tribe, Harry].

Harry may be willing, the tribe may be, or if they're in luck, both will be, but 18) doesn't clearly express any of these. If both parties are willing, for example, we need two occurrences of willingly, as in

(19) Willingly(λv_1)willingly(λv_2)(v_1 sacrifices v_2)[the tribe, Harry]

(where one occurrence of a willingly to the right of the other here does not affect its contribution - its not that, roughly, the tribe is only willing to sacrifice a willing Harry.) As Davidson almost said, intentionality is tied to a person or a tribe. In "intentional actions are not a

class of actions"(cp. [14], pp. 244, 245), we have an intimation of something remarkably similar to the distinction proposed above. It is: mode of doing an action, versus type of action done. The difference is marked by an asymmetry. We can say that Jones deliberately buttered the toast slowly, meaning that he took pains to do it slowly. But "Jones slowly buttered the toast deliberately" can only mean that he did it slowly and he did it deliberately. A deliberate buttering is not a type of action, something which is open to qualification according to speed, in the way that a slow buttering is open to qualification with regard to intention.

One cannot straightforwardly read off 19) in ordinary English and, by the same token, it may be complained against 17) that necessarily does not occur twice in Hesperus is necessarily Phosphorus. But given the special prominence of the grammatical subject, it may be that that sentence is best represented by NEC(λv)(v is Phosphorus) [Hesperus]. This is not to deny that there are more blatantly symmetrical cases, eg. Hesperus and Phosphorus are necessarily identical - which, it may be argued, involve a genuine relation of necessary identity. But I would still maintain that 17) expresses an essentialist thought: that both Hesperus and Phosphorus necessarily-partake of identity with one another. Neither of i) and ii) is well-suited to NEC, the essentialist's copula modifier, so we could take option iii), treating it as forming terms from 1-place predicates. Semantically, the natural treatment would be just like that of 1) for plain λ , ie. with "NEC λv_i " replacing " λv_i ", and "NEC λv_i " " λv_i ". Multiple NEC abstracts, as in 17), can obviously be treated parallel to 2).

We may envisage a class of term-forming predicate modifiers, of which plain λ is the unadorned limiting case. (For notational continuity I have included λ in $(NEC\lambda v_i)$, though really it is neither syntactically nor semantically a component.) Whether treatment along the lines of 1) is ultimately successful depends, as the Wallace argument serves to remind us, on whether the modification of the satisfaction predicate which results in the processing of a sentence containing the modifier mirrors the modification that occurs in that sentence. It would thus fail for the blatantly intensional willingly. What does it take to willingly satisfy assassinate Maggie? How can you noisily satisfy eat an apple? I imagine that satisfaction, being an abstract relation, is a rather quiet affair. Whereas modification of T_{sc} by modalities and tenses can be agreed to be intelligible, with this adverb it is not so clear⁴. But we have still to face the question of whether NEC and Nec modification of T_{sc} , on the RHS, matches what occurs on the LHS.

The following suggests itself by analogy with axiom Sat 3 of Wiggins' miniature theory for Nec ([109], p.305):

$$(20) \text{ For all } s \ (T_{sc}(\text{Nec}(\lambda v_2)\text{Man}(v_2)[\text{Caesar}], s, F_4) \equiv \text{Nec}(\lambda v_2) T_{sc}(\text{Man}(v_2), s(2/v_2), F_4)[R_{sc}(\text{Caesar}, s, F_4)]).$$

What the RHS of 20) says - or better, its NEC counterpart - is that the referent of Caesar, whatever that is (it matters not how it is described), necessarily has the property of being the second member of a sequence satisfying in F_4 Man(v_2): It is not claimed that "the English expression Man prefixed to the variable v_2 is such that it must be satisfied by any sequence with Caesar as its second member ... [The kind of claim made] is that certain individuals (via

sequences) necessarily-satisfy certain expressions; not that these expressions are necessarily-satisfied-by those individuals" ([109], p.311). But does this really help? The idea is that, simplifying out the sequences, we don't have this de re attribution:

(21) $(NEC\lambda v)(v \text{ is true-in-} F_4 \text{ of Caesar})[\text{Man}(v)]$

but rather this:

(22) $(NEC\lambda v)(\text{Man}(v) \text{ is true-in-} F_4 \text{ of } v)[\text{Caesar}]$.

It follows from what has been observed in connection with 11) that an existential asymmetry arises between a term α and any term in a formula φ in an essentialist claim of the form eg. $(NEC\lambda v)(\varphi(v))[\alpha]$. Its reformulation in terms of the de dicto ($\alpha \text{ exists} \rightarrow \varphi(\alpha)$) brings out clearly the existential caveat on α (cp. [79], p.322, and [109], pp.310-311). So in 22) we have a claim which entails that Caesar would not exist in any situation in which either the expression or the language failed to exist. Moreover, when we come to consider the proper version in full we shall not only find mention of a sequence of objects but also the long suppressed parameters of utterer and time. Similar problems beset the embedding of truth theoretic apparatus within \square , for any fragment containing demonstratives and quantifiers. For \square we could avoid these problems by restricting the concern of truth theory to weak necessity, a stratagem which by definition fails to help with strong necessity, Nec and NEC.

One might appeal at this point to the abstract nature of some of the entities involved - language fragment, expression, sequence and time. The alleged difficulties with the dependence of Caesar on them effectively disappear if these are entities which cannot fail to exist. But being an abstract object is not in itself a passport to every

possible world, as Wiggins' discussion of the set {Crystal Palace, The Eiffel Tower} testifies ([110], section 4.4). Since a set is defined in terms of its members, its existence depends on theirs (though not conversely). A set such as this one whose members are contingent beings itself exists contingently: if one or other of its members had not existed, it would not have. Nevertheless it might be argued that the abstract entities we are concerned with need not be regarded as contingent existants. This suggests that one may go some way to fend off the problems which arise through the embedding of truth theoretic apparatus within modal operators if one is prepared to pay a certain price. A similar conclusion is suggested by the response to the Wallace argument which was shelved earlier. Consider first the argument as it applied to the use of the past tense operator in the ML. The argument was that eg. something of the form $PT(\varphi, \text{English})$ could be true even though $T(P\varphi, \text{English})$ is not, because some term in φ had changed its meaning. But when the predominant meaning of a word does change - gay being an example which readily springs to mind - its not as if the old meaning is totally erased from the language. Such a case should therefore be treated as a normal case of ambiguity (section 1.1), and once disambiguated, the argument does not go through. But if there is a word for which, if you go back far enough, you can find a meaning which has no currency at all today (as with villain, perhaps), then you may legitimately decide to discount that as a part of the English language investigated as the CL of your truth theory. That is a plausible delimitation of what actually counts as the language or fragment under study (the main cost of which being to regard the line between

ambiguity and antiquity as sharply drawn). But it therefore fails to relate to the question of what could or must count as English. To say that we identify English as the language in which, inter alia, is is the copula⁵, is not yet to say that it could not be, in this language, eg. a primitive sign for distinctness rather than identity. So the price of the response to the Wallace argument is higher in the modal case than it is with temporal operators. Is it worth paying? Ultimately, the worth of the kind of theory we have been considering in this section depends on its incorporability into a theory for the whole of English. And there is a special problem which arises for any theory coping with a fragment containing both modal and indexical locutions, as the following application of 1.2.13 illustrates:

$$(23) T_{sc}(\Box \text{man}(\text{that}), s, u, t, F_4) \equiv \Box T_{sc}(\text{man}(\text{that}), s, u, t, F_4).$$

The LHS of 23), we may assume, is true - if u demonstrated a man at t. The RHS, however, says that it is necessary that man(that) is true, relative to the given parameters - something arguably false, given that u might have pointed to a frog. The falsity is underlined by taking a further step using the Davidsonian axiom 1.2.3, to $\Box \text{man}(\lambda v \text{Dem}(v, 1, u, t))$. Demonstratives are both essential to a natural language and such that their reference, on any given occasion of use, is a contingent affair. But perhaps all this point proves is that all such statements must be construed as judgements de re. There is no problematic correlate of 23) if the claim is formulated as $\lambda v \Box \text{man}(v)[\text{that}]$, for example. (This is not to say that there could not be de dicto necessities involving indexicals. In this sense, perhaps, it is necessary that I am where I am, you were born where you were born, he is who he is, and so on - but such cases cannot be

problematic in the way 23) is.)

5: MODALITY AS QUANTIFICATION

In this section I want to consider a few of the many issues that are raised by the interpretation of modal locutions using quantification over such things as possible worlds. A good place to start is with the argument of Davidson and Wallace to show that this style of interpretation should be homophonic, ie. should read the world variables into the expressions of the OL. In Davidson's version, we start with a thesis proposed by Tarski, that natural languages are in essence intertranslatable. Any natural language should be able to express, whether by means of a simple or a complex expression, what is expressed by any word from any other such language (cp. [19], p.82).

Wallace puts forward a more precise principle: "if S_1, \dots, S_n are sentences of L_2 [eg. the ML] that translate some sentences of L_1 [eg. the OL] and if S_{n+1} is a sentence of L_2 built from vocabulary that occurs in S_1, \dots, S_n then S_{n+1} translates some sentence of L_1 " ([106], p.225). We can see these principles at work by considering the modal case.

Suppose we interpret sentences of the form $\Box\varphi$ from L_1 using those of the form $(\forall w)\varphi'(w)$ in L_2 . Then from the vocabulary of the latter we can construct sentences about the existence of exactly n worlds such that ..., and so on (assuming L_2 has identity), which are inexpressible in the terms of the sentences with simple operators. Wallace's principle counsels against such an account of form and scheme of interpretation built upon it. But suppose L_1 has the adjective possible and the means to express what a philosopher intends by the world, ie. "everything that is the case" - as we would expect of a natural language - then there will be

nothing untoward in the relation described between L_1 and L_2 . Both principles are agreed that the interpreting ML should be translatable back into the OL. And so, Davidson's line of argument runs, since the theory of L_1 in L_2 teaches us that the use of operators is really possible worlds talk (for that's how they're interpreted in the ML), and since this is the expressively more fundamental idiom (and intelligible to the OL speakers anyway), we should take it that really (at a level beyond the surface) the OL sentences involve the possible worlds idiom also¹.

Perhaps the simplest homophonic approach just requires the interpretation, in a theory like Θ_2 , of three special predicates: $PW(v)$, $PRT(v_1, v_2)$ (for v_2 's being possible relative to v_1), $In(v_1, v_2)$ (for v_1 's being in v_2 - a world-relative existence predicate, in effect - to cover the interaction of modality and quantification), and the special constant @. A second kind of homophonic theory might use the world-relativised truth predicate T_{sw} typically employed by a heterophonic theory such as Θ_6 . A crucial axiom for Θ_6 is this:

$$(1) \quad T_{sw}(\Box\varphi, s, w_1, F_3) \equiv (\forall w_2)(PRT(w_1, w_2) \rightarrow T_{sw}(\varphi, s, w_2, F_3))$$

(once again u and t are suppressed). But the significance of 1) cannot be judged on its own. It must be taken in conjunction with clauses for the basic predicates

$$(2) \quad T_{sw}(\alpha \text{ is brave}, s, w, F_3) \equiv R_{sw}(\alpha, s, F_3) \text{ is brave-in } w$$

a clause for the quantifier, eg.

$$(3) \quad T_{sw}(\exists v_i \varphi, s, w, F_3) \equiv (\exists y_i) In(y_i, w) \& T_{sw}(\varphi, s(i/y_i), w, F_3)$$

a clause for actuality:

$$(4) \quad T_{sw}(\underline{\Delta}\varphi, s, w, F_3) \equiv T_{sw}(\varphi, s, @, F_3)$$

and one to remove the relativity to worlds from truth:

$$(5) \quad T(\varphi, F_3) \equiv Closed(\varphi) \& (\forall s) T_{sw}(\varphi, s, @, F_3).$$

We could have relativised the notion of reference to possible worlds, and indeed one would do so if one thought that the reference of some genuine referring expressions was sensitive to that parameter. Evans counsels against such a measure ([39], p.169). He also reminds us that " $T_{sw}(\varphi, s, w, F)$ " is to be understood as satisfaction with respect to the world w - it is a world of evaluation, not utterance (cp. Gupta's proposal, §1.3).

At least one major disadvantage of a homophonic theory, unlike a theory such as Θ_6 , is that it requires all the predicates of the CL to be relativised to a world variable - even when modality is not the theme. But against this charge of blatant violation of CL syntax, Davidson and Wallace would presumably counter that the heterphonic Θ_6 anyway interprets apparently n -place predicates as $n+1$ -place, in the ML - as 2) makes apparent. The basic problem, it seems to me, affects both kinds of theory, and is this: most of the time possible worlds do not intrude into our everyday talk, and even when they do (eg. with "We could have had a situation in which ..."), it is implausible to say that we are dealing with a primitively world-relativised predicate. These considerations suggest that we might consider the possibility of introducing what might be termed an 'indexing operator' \mathbb{E} , an expression which operates on open or closed sentences to produce something of the same ilk, but with one more argument place. $\mathbb{E}(w)\varphi$ should be understood as: φ with respect to w . Let us try to construct from Θ_2 a theory Θ_7 for \mathbb{E} along the following lines. Θ_7 is to employ odd-indexed variables v_i to range over ordinary (possible) individuals, even-indexed v_i to range over worlds (ie. we assume that these are restricted by the predicate PW

and can also be written w_1 .) Quantification is interpreted as before (axiom 1.2.10). The basic idea for \mathbb{E} is to use it in the ML to modify the truth predicate when needed:

$$(6) \quad T_{sc}(\mathbb{E}(\beta)\varphi, s, F_3) \equiv \mathbb{E}(R_{sc}(\beta, s, F_3))T_{sc}(\varphi, s, F_3).$$

Actuality will now be treated with a reference clause:

$$(7) \quad R_{sc}(@, s, F_3) = @$$

and we shall need the familiar-looking rule:

$$(R\mathbb{E}) \quad \text{If } \vdash_{e_7} (\forall v_1) \dots (\forall v_n) (\varphi \equiv \psi) \text{ and } \vdash \mathbb{E}(\beta)\varphi \vdash, \\ \text{infer } \vdash \mathbb{E}(\beta)\psi \vdash.$$

For example, if we take the sentence

$$(8) \quad \text{There could have been things other than those which} \\ \text{actually do exist,}$$

representing it thus:

$$(9) \quad (\exists w_2) \mathbb{E}(w_2)((\exists v_1) \mathbb{E}(@)((\forall v_3) v_1 \neq v_3))$$

we find that

$$(10) \quad T_{sc}((9), s, F_3) \equiv (\exists w_2) \mathbb{E}(R_{sc}(w_2, s(2/w_2), F_3)) T_{sc}(\underbrace{(\exists v_1) \mathbb{E}(@)}_{((\forall v_3) v_1 \neq v_3)}, s(2/w_2), F_3).$$

A repeated application of 1.2.10 and 6), the equations $R_{sc}(w_2, s(2/w_2), F_3) = w_2$ and 7), some more of e_2 and a couple of applications of $R\mathbb{E}$ will see us through to the desired truth condition. It is perhaps worth pointing out that e_7 can be extended to cover other intensional constructions in the following way:

$$(11) \quad T_{sc}(\alpha \text{ believes that } \varphi, s, F_3) \equiv (\forall w_2) (\text{Belief}(R_{sc}(\alpha, s, F_3), \\ w_2) \rightarrow \mathbb{E}(w_2)T_{sc}(\varphi, s, F_3)).$$

Since we are reading \mathbb{E} as "with respect to", we avoid the problematic ascription to Maggie et. al. of truth-theoretic beliefs, desires and so on. In resolving propositional attitudes thus the theory buys off one set of unwanted inferences but is still left with another set, slightly more palatable perhaps, but still undesirable (see below, §2.4).

As an alternative to Θ_7 - if, for example, we do not care for the use of "with respect to" to modify predicates in the OL - we might attempt to make a theory Θ_8 for "in", symbolised $\$$. Since we ought not to allow truth-theoretic material to occur within the scope of $\$$, we should aim to design Θ_8 around the primitively world-relative truth predicate T_{sw} of Θ_6 . The simplest way would be, keeping $\$$ out of the OL, with a theory like Θ_6 except with regard to the basic predicates:

$$(12) T_{sw}(\text{brave } \alpha, s, w, F_3) \equiv \$ (w) \text{ brave } R_{sw}(\alpha, s, F_3).$$

It is only here that $\$$ is really vital, since "PRT" and "In" will cover the other interactions between a modality and an expression. Adopting a rule to allow φ to be inferred from $\$(@)\varphi$ will allow us drop reference to worlds (ie. the actual one) when modality does not enter into the conversation.

But to get $\$$ into the OL we need an axiom like this

$$(13) T_{sw}(\$ (\beta)\varphi, s, w, F_3) \equiv T_{sw}(\varphi, s, R_{sw}(\beta, s, F_3), F_3).$$

β here will either be $@$ - in which case see 7) - or a world variable w_i - in which case it is basically treated along the lines of 1.2.8 (by alternating the indices of ordinary and world variables, there will be no conflict between individuals and worlds for the places of a sequence.)

Complementing 13) is the axiom for world quantification:

$$(14) T_{sw}(\exists w_i \varphi, s, w, F_3) \equiv (\exists w_i) T_{sw}(\varphi, s(i/w_i), w, F_3).$$

In assessing the relative strength of the $\$$ against the $\$$, it is worth considering the expressive resources deployed in a sentence like 8). 8) cannot be represented in a language containing only the operators \Box and \Diamond . The problem for the simple operator language which this sentence presents is that initially we need the combination $\Diamond(\exists v_1)$ to assert the possible existence of something, but then to

follow it with $(\forall v_3)$, within the scope of the quantifier but not within the scope of the possibility operator (cp.[50], p.34). One needs to introduce a 'backwards-looking operator' A (cp. 4) above) which, as now always does with the present moment of time in temporal contexts, refers us back to the actual world, no matter what modal operators it occurs within the scope of. Actually, as it is actually used in English, is a device of contrast and emphasis - as for example when we say that theoretically such and such, but actually so and so. It does not seem to have the semantically significant function performed by A, which is rather carried out by the mood of the verb. According to Lewis, actually is ambiguous: in its primary sense, it refers us back to the actual world (what is for him a world of utterance), whereas in its secondary sense it refers us to some other world which has been introduced for consideration. The phenomenon he is aiming to account for is manifested by the following pair of conditionals:

(15) If Max ate less, he would be thinner than he actually is

(16) If Max ate less, he would actually be thinner than he is

But the alleged ambiguity would be better explained by a redundancy theory of actually, according to which it adds nothing (but emphasis) to what is achieved by the mood of the verb. If actually adds nothing to the cognitive content of the conditionals, we should be able to remove it without affecting their sense. This is indeed so:

(17) If Max ate less, he would be thinner than he is.

Similarly, I have been unable to find a way of formulating

8) in which actually clearly plays an essential role, ie.

which prevents it from lapsing into contradiction. It is

words like do (and even exist on its own) and is, in 8) and

15), which achieve the effect of referring us back to the actual world and which therefore demand the presence of A in their representation. (The adjective actual is more active, however - cp.2.4.5.)

If we give the necessity and possibility operators indices: \Box_i , and \Diamond_i , we can introduce a family of these backward looking operators A_i , any of which will refer us back outside the scope of an intervening modality to the prior occurrence of a modal operator with the same subscript (cp. [80]). In $\Box_i \dots A_i$ we have a notational variant of eg. $(\forall w_i) \dots \S(w_i)$, in so far as the cross-referential effect of both systems is the same. Each of the A_i , $\S(w_i)$ and $\S(w_i)$ notations have the effect, when added to a language, of converting it into a first order language whose predicates have an argument place for world variables and @. But there is a difference. There is an obvious difference between the actuality operator A and the term @ (or between now as operator and as indexical constant), in that the latter can stand as the subject of predication, identity and so on. In a sense, this is a theoretically dissoluble distinction. Quine's "Variables Explained Away" shows how an operator language can be a notational variant of a first order language. Suppose we have the operators \Box , \Diamond and A, and \forall and \exists for the standard quantifiers. We adopt the convention that an operator binds the last available argument place of a predicate. To express all possible predicate-logical formulae we need other operators to juggle the order of argument places of a predicate. For example, if loves is the familiar predicate whose second place is reserved for the object of the verb, $\exists(\text{loves})$ is the predicate, in quantifier/variable notation, $\lambda x(\exists y)(x \text{ loves } y)$.

If INV is the operation which, for a two-place relation produces its converse, ie. switches the order of the arguments, then binding the second place of the modified relation, $\exists(\text{INV}(\text{loves}))$, is the same as binding the first place on the old predicate: $\lambda x(\exists y)(x \text{ is loved by } y)$ ($= \lambda y(\exists x)(x \text{ loves } y)$). Other operators will be needed to handle predicates of higher^{arity} adicity. So with the actuality operator, $A(\text{loves})$ would express the property of loving the actual world $\lambda x(x \text{ loves } @)$, a peculiarly Panglossian sentiment. We could of course introduce another operator $\$$ such that $A(\$(\text{loves}))$ expresses $\lambda x \lambda y(x \text{ loves } y \text{ in } @)$. This shows up how Quine's important insight into how an operator language can be extended to a quantificational one (cp. van Benthem's detailed survey in connection with temporal discourse, [104]) remains a theoretical one as far as we are concerned. For clearly our ordinary modal operators do not follow the convention of binding the last or any other argument place of a predicate - witness the fact that we do not require the insertion of $\$$ to make intelligible the ordinary use of It is possible that. And it brings me back to the difference between the notational systems we have been dealing with, which concerns the argument places which $@$ and w_i are allowed to occupy. If we are to say that the $\Box_i \dots A_i$ system, when added to a first order language, produces just a notational variant of a two-sorted language, it is one in which the two sorts of variable are kept clearly segregated. The second sort of variable must be regarded as inferior in that their role is restricted to indexing assertions about the first sort of entity over which the ordinary variables range - it allows for no judgements solely about worlds.

In concluding this brief section I would like to say

that perhaps too much is made of the lack of explicit reference to worlds in average English modal discourse. For we have seen that there are various ways of extending our ordinary operator language (of \Box , \Diamond and A , say) into a quantificational one, not the least of which involves mixing in an "in" operator $\$$ and the relevant variables. It seems to me that if ordinary speakers had an interest in expressing modal distinctions more finely they would resort to the extensional idiom. The moral here is drawn from the temporal case: we do have a need to go beyond what Geach calls 'grass-roots' temporal discourse, involving only the operators F and P - which on their own merely act as pointers, pointing either forwards or backwards in time - to express distinctions which require the explicit mention of moments or intervals of time. In this respect it is significant both that the possible worlds idiom is not hard for an ordinary speaker to understand and that philosophers - who by tradition have more interest than most in getting modal issues straight - will resort to the idiom when convenient, even if they prefer not to countenance worlds as genuine members of their ontology.

CHAPTER TWO: MONTAGUE AND MODEL THEORY

1: STRUCTURAL SEMANTICS AND MODEL THEORY

In this section I shall attempt to elucidate the nature of the semantic programme largely inspired by the writings of Richard Montague. I start with the distinction between lexical and structural semantics. The distinction is natural, and easy to state in general terms. Lexical semantics is concerned with the specific meanings of the various words which occur in some language. Structural semantics is concerned with the way these meanings can combine to form the infinity of different meanings expressed by the infinity of different sentences of the language.

Structural semantics deals with the forms of contribution to truth conditions (or whatever the central semantic property is taken to be), lexical semantics deals with particular contributions. Content and form, base and recursion - what could be simpler?

To make the distinction is not by itself to make a commitment to the view that they are separable enterprises. Lexical semantics, if it is to systematically comply with Frege's admonition only to ask for the meaning of a word in the context of a proposition, must presuppose the general insights of structural semantics in order to give the proper forms to its meaning specifying clauses. On the Davidsonian conception of a theory of meaning, the main function of structural semantics is to provide answers to the question: What is the form of clause for this kind of word (or construction)? On that conception, the two components of

semantic theory are inseparable parts of a single empirical enterprise, in that they would not be tested independently but through the products of their interaction, the T-sentences. Nevertheless, the structural component must have a degree of independence from the lexical. One can pose the question: what is the mode of contribution of this type of word? without inquiring too far into the particular specimens which function in that way. This becomes clearer from consideration of Montague's chosen method for the semantic investigation of a language, model theory, where the particular meanings of most expressions are left schematic.

In extensional model theory, EMT, it is customary to take a model \mathcal{M} for a language L to be a pair $\langle A, F \rangle$. A is a set of entities - as before (81.2), the domain of individuals talked about in L . F is a function, an assignment of appropriate extensions to the constants of L . In particular,

(1) For any name $\alpha \in L$, $F(\alpha) \in A$

- names denote entities. To a 1-place predicate φ it assigns a set, which can be regarded as a function from entities to truth values, the characteristic function of the set. Two members of A may be conventionally chosen as the truth values truth and falsity, eg. the numbers 1 and 0 respectively. So

(2) For any 1-place predicate φ , $F(\varphi) \in \{1, 0\}^A$.

The semantic classification of expressions is effected through the types of their possible denotations. The basic types are e , that of entities, ie. members of A , and t , that of the two truth values. The set of types is defined as the smallest set Y such that i) $e, t \in Y$ and ii) if $a, b \in Y$, $\langle a, b \rangle \in Y$. $\langle a, b \rangle$ is the type of functions from things of type a

to things of type b . Thus the type of a name α is e , and of a 1-place predicate φ is $\langle e, t \rangle$.

F enables the model theorist to by-pass the truth theorist's lexical-semantical lists of clauses for the non-logical constants (cp. 1.2.2). It allows him to characterise the extensions of these expressions in general, semantical terms. The reference of names α is characterised, relative to a model $\mathcal{M} = \langle A, F \rangle$ and a value assignment g of values (in A) to variables, by the one clause

$$(3) \quad R_{ms}(\alpha, g, \mathcal{M}) = F(\alpha)$$

(relativity to a language is implicit in \mathcal{M} , a model for L , whose syntax should really be given.) For variables β

$$(4) \quad R_{ms}(\beta, g, \mathcal{M}) = g(\beta).$$

Satisfaction in a model, T_{ms} , for 1-place predicates φ gets the clause

$$(5) \quad T_{ms}(\varphi(\alpha), g, \mathcal{M}) \equiv R_{ms}(\alpha, g, \mathcal{M}) \in F(\varphi).$$

The clauses for and and not are generalised for \mathcal{M} in the obvious way, and if β is a variable of type e ,

$$(6) \quad T_{ms}((\exists \beta)\varphi, g, \mathcal{M}) \equiv (\exists v \in A) T_{ms}(\varphi, g', \mathcal{M}),$$

where g' is like g except, if at all, in assigning v to β - is the quantifier clause. Truth in a model, $T_m(\varphi, \mathcal{M})$, is defined parallel to 1.2.11, ie. in terms of satisfaction by all valuations. Truth in a domain, $T_d(\varphi, A)$, is truth, no matter how the constants are assigned extensions (in A), ie. $(\forall F) T_m(\varphi, \langle A, F \rangle)$. Removing this final relativity brings truth in every (non-empty) domain, $T_1(\varphi)$, which is $(\forall A)(\forall F) T_m(\varphi, \langle A, F \rangle)$, ie. logical truth.

In order to get at what the model theorist accomplishes with this truth definition, it is helpful to compare it with the correspondence theory. The theories share two fundamental characteristics. i) "It makes truth a relation".

This - usually uttered accusingly - needs to be qualified. As far as the correspondence principle is concerned - a true sentence/statement is one which corresponds to the facts - it is correspondence which is the relation, the obtaining of which is concurrent with the holding of the unary property of truth. Secondly, even a Tarskian 'absolute' theory makes sequences a crucial relatum of truth. Indeed, Davidson has argued that this kind of theory deserves to be deemed a correspondence theory, precisely because it explains truth in terms of a relation between words and world, satisfaction ([16], p.758). It is tempting to think that the Tarskian theory in fact explains truth in terms of two relations, not one: satisfaction and reference (roughly, the descriptive and demonstrative relations with which Austin sought to account for correspondence.) But if we believe that the sentence/statement is the basic unit of currency in linguistic transactions, this must be the wrong way of explaining things - it puts the small change first, rather than the basic property, truth. There is a circle here - as Dummett puts it: "in the order of explanation the sense of a sentence is primary, but in the order of recognition the sense of a word is primary" ([29], p.4). (For this reason, one can accept that truth is basic without thinking that evidence for a semantic theory cannot accrue at the sub-sentential level.) Or, in terms of the correspondence theory, since there is no independent account of what a fact is, there is no explaining the ordinary, unrelativised concept of truth by means of the relativised concept of correspondence.

ii) Both theories are, in an important sense, schematic. Understood as a straightforward piece of English

the correspondence principle conveys two truisms: that factual, ie. capable-of-being-true, discourse is factual, ie. concerning-itself-with-facts, discourse; and that that which is true is that which does correspond to the facts. If we do not consider a statement to be of a kind which states facts, we do not consider it a fit subject for assessment as to truth or falsity. The principle is thus a proto semantic theory of the nature of truth: a true statement is one which has this crudely articulated semantic function: it corresponds to the facts. The principle itself says nothing about which sectors of discourse are to count as factual, ie. what would count as a fact, and what would not. (Thus with the additional assumption which Austin made, that it is "simply not the business of such utterances [as eg. of arithmetical formulae] to 'correspond to the facts'", he drew the conclusion that it would be wrong to evaluate them as true or false ([2], pp. 131-32). Similarly, the definition of truth provided by BMT doesn't by itself rule on what kind of sentence is to count as true or false. Consider for example subjectivity, a potential blemish of certain natural language predicates. If beauty really is in the eye of the beholder, it cannot faithfully represent the nature of the predicate beautiful to say that there is a set of all and only the beautiful things, uniquely determined as its extension - as 5) in effect does. If two people disagree over whether something is beautiful, it is not simply that they have different beliefs; if beauty genuinely is a subjective matter, there is no 'in principle' correct answer to whether something falls within the extension of beautiful - even if there is agreement amongst speakers on the intension. The question is not whether

EMI rules out beautiful, but whether there is no way of extending it to provide a definition of something like truth for subjective predicates.)

The correspondence principle is sometimes put in terms of the 'situation formula' : a statement is true if the situation or state of affairs stated to obtain does obtain. We can reformulate the latter for the simplest case of a statement of subject-predicate form as: if the thing named as the subject has the property expressed by the predicate, or less richly: if it is one of the things the predicate applies to (cp.[46]p.126). In spelling out the idea of correspondence in a little more detail, by thus giving one part of what would have to be a recursive definition for the different sentential forms of the language¹, we have something recognizably an informal counterpart to the RHS of the model theorist's clause 5). If the parallel is correct the LHS of 5), the truth of $\varphi(\alpha)$ in a model is the formal correlate of $\varphi(\alpha)$'s corresponding to the facts.

One may conceive of changing the assignment F in two ways, one linguistic, the other factual. The fact that there will be F_1, F_2 such that $F_1(\alpha) \in F_1(\varphi)$ and $F_2(\alpha) \notin F_2(\varphi)$ may be conceived of as a variation in the meanings of the expressions, and therefore of what they have as extensions, the facts remaining the same, or as a variation in the facts of the matter, ie. what falls within what set, the meanings held constant. On the former conception logical truth, T_1 , truth no matter what assignment F and domain are chosen, comes out as truth, no matter what the non-logical constants mean or, if you will, no matter what language - of this structure - you chose. On the other, logical truth is the familiar idea of truth in all possible worlds. The

important thing is that T_1 is truth solely in virtue of composition (in the right way) out of the logical constants. Actually, there is a slight problem with the second conception. There is nothing to stop there being an assignment F which includes the denotation of Caesar amongst the things which fall under frog, or which assigns circular and square overlapping extensions. This is the familiar point, pressed eg. by Quine against Carnap ([88], p.23), that logical truth, in the wide sense, is wider than metaphysical truth in the not so wide sense. And not only are models dubious surrogates for possible states of affairs in allowing in more than is genuinely possible, we would also expect there to be differences between states of affairs which models cannot match, since the latter are relative to the expressive resources of a particular language (cp. [73], pp.153-54). Because the parallel between models and possible facts is therefore not exact, it would be best to say that the relative notion of truth in a model explicates the idea of correspondence with the facts. In so doing, the model theorist shows that the wrong things were expected of the schematic idea of correspondence - its use is not in the explanation of ordinary, unrelativised truth but, by blotting out the particular meanings of the non-logical constants, to enable a definition of T_1 .

If the model theorist is to bear out a claim to be providing structural semantics, he must explain why this discipline should be so interested in the meanings of one group of words, the logical constants - or, to put it another way, why a theory of semantic structure should provide a theory of logical form. Logical form is a notion which tends to get overworked, but its basis must surely be

the idea of the form a sentence has which is relevant to its role in logical arguments. So a theory of logical form, for the sentences of some language, will presuppose an account of logical argument - a characterisation (whether syntactic or semantic) of consequence. A set of logical expressions will be distinguished from the nonlogical, and it will be their interpretation which counts, as far as the logical form of a sentence goes, the rest being irrelevant. Thus on the normal reading of

(7) Something loves everything,

of the various obvious candidates a) $(\exists v_1)(\forall v_2)\text{loves}(v_1, v_2)$, b) $(\exists v_1)(\forall v_2)\Psi(v_1, v_2)$, and c) $(Q_1 v_1)(Q_2 v_2)\Psi(v_1, v_2)$, it is representation b) which optimally reveals its logical form. The logical form of a sentence presupposes an interpretation of the logical constants, as Grandy has emphasised ([45], p.157), and strictly speaking, no more. As the example also shows, a theory of logical form for a natural language will also require a mapping, specifiable in principle, from its sentences onto their forms. This mapping will enable us to produce quasi-English sentences such as a), which reveals the logical form of 7) in virtue of being a paraphrase of it in accordance with some theory of logical form.

The natural way to understand semantic structure is as semantically relevant form. If structural semantics is concerned with the ways in which the words of a language can contribute to the meanings of sentences in which they occur, it must produce groupings of words according to their shared semantic function. Since the semantic structure of a sentence will not concern any particular contribution to its truth conditions, it would best be displayed by showing the sentence as a collection of members of such and such

semantic categories in such and such a combination. 7) is made up by a combination of two quantifiers and a two-place relation; c) is therefore the representation which optimally displays its semantic structure. a) and b) (in conjunction with an assignment of the words to their semantic categories) also reveal the semantic structure of 7), but they do more besides. So why equate semantic structure (c)) with logical form (b))?

Structural semantics, on the conception outlined so far, will be based upon very general categories of words - giving the common semantic function of nouns, of adjectives, and so on. The pressure for extending this basis to provide more detail seems to stem from a desire to account for certain inferences as valid in virtue of the semantic composition of the sentences involved. The more general the semantic categories, the sparser the collection of structurally valid inferences. If we lump all and only the binary truth-functional connectives into one semantic category C, an inference concerning and will count as structurally valid just in case it depends on the truth-functionality, not the particular truth function, of the word, ie. if it is an instance of substitution of material equivalents. But C does not correspond to the natural syntactic category of binary connective, since there are plenty of non-truth-functional connectives. The more inclusive category would produce an even weaker characterisation of structurally valid inference. Similarly, a general characterisation of adjectives β will discount an inference α is a $\beta\delta$ to α is a δ as structurally valid, in view of such adjectives as possible ([73], p.211). So instead one might approach the notion from a different

direction, as Davidson does. He holds that the inference from "Jones buttered the toast in the bathroom" to "Jones buttered the toast" is a structural one, in view of his intuition that "it has something to do with the fact that 'buttered' turns up in both sentences" ([18], p. 143). This imparts a "common conceptual element" to the two sentences. The entailment is structural because "it depends on the fact that the word 'buttered' is playing a certain common role in both sentences" ([18], p. 144), i.e. the way it contributes to the truth conditions of the whole is the same in both cases.

The notion of shared function in conclusion and premise cannot be enough by itself to make a criterion of the structural inference. All the words in a sentence φ contribute to its conditions of truth in the same way that they do in Not φ (as on Davidson's paratactic theory, they should also do in Jones believes that φ , etc.) - yet Not φ does not entail φ on any conception of inference. The obvious extra ingredient is that in addition the conclusion must be a logical consequence of the premise. An interesting demarcation of structurally valid inference will require a theory of semantic structure to take account of the particular meanings of the logical constants. A second point to note is that the fact that "buttered" plays a common role in "Jones buttered the toast in his bathroom" and "Jones buttered the toast" depends primarily on the way "in his bathroom" modifies the latter. To adapt an example of Montague's, the inference fails if we take "Jones buttered the toast in his dream" as the premise ([73], p. 213). If Davidson's inference is structurally valid, it must be because the basic category of adverbial phrase has some significant subcategories. If the general structural

semantic characterisation of that category is adequate, it will be capable of spelling out the common function of words in these subcategories as special cases, and thus explaining their special behaviour. This is in practice what will happen on a sufficiently rich model theoretic account. For example, given the simple theory of adjectives as applying to the intensions of common nouns, the properties they express, adjectives β which permit the inference from α is a $\beta\delta$ to α is a δ are those which produce a subproperty of the property expressed by the common noun δ ([53], pp.124-125).

On the present conception, structurally valid inference is not to contrast with logically valid, which it must include, but with what might be termed lexically valid inference - that which depends on the particular contributions of certain elements. The inference from Jones knows that φ to φ will count as structurally valid if either know is a logical constant, or if a suitable subcategory is identified, of factive sentential adverbial phrases, otherwise it will count as lexically valid. This strongly suggests that this conception has an obligation to provide us with some demarcation of the logical constants, since it deliberately permits the specific lexical contributions of the logical constants to figure in structurally valid inference.

Davidson has offered a criterion of logical constancy: treatment by a recursive clause in the definition of truth. The criterion is, however, notoriously partisan. It picks out the standard logical constants from a standard first order language - excepting identity - but it is possible to treat blatantly non-logical words with a recursive clause: at least the extensional adjectives such as four-legged,

treated as predicate modifiers within the context of a Tarskian 'absolute' truth definition (as Evans in effect pointed out in [37], p.203).

The recursive clause idea can be made into a working criterion if it is formulated within the context of a model theoretic definition of truth. An expression is treated as a logical constant by a theory if it is singled out for attention by having its meaning specifically stated. The non-logical words are those whose meaning is given schematically, i.e. through a basis clause involving F or g . The working model theorist, after all, is not confined by some philosophical criterion of logical constancy, but is concerned with the syntactic and semantic characterisation of some expression or construction. The description of the semantic function of non-logical words in general terms goes hand in hand with the characterisation of other expressions in detail. A truth definition which captures the indefinite contribution to truth conditions of vague predicates will provide the means for a definition of an operator "It is definite that"; a description of the intensions of words, understood as functions from possible worlds to extensions, will enable one to define \Box and \Diamond , and so on.

EMT characterises the truth and falsity of sentences purely in terms of extensional features of their parts. Frege would have said that it characterises the extensions of sentences in terms of the extensions of their parts. In compliance with his view we could have written $R_{ms}(\varphi, g, m) = 1$ instead of $T_{ms}(\varphi, g, m)$. He would have added that it was therefore inadequate to deal with certain constructions. "It is a priori that Hesperus is Hesperus" is true, but that cannot be a function of the extensions of the parts, for the

parts of "It is a priori that Hesperus is Phosphorus" share those extensions, but it is false. Frege put this by saying that in such contexts as "It is a priori that" a sentence refers to its sense - translating out of his terminology, we may say that the characterisation of the truth and falsity of sentences will in general require us to consider the senses or intensions of their components. Montague's intensional model theory of FFO, IMT, achieves just that.

An intensional model, or interpretation \mathcal{M} for the language of tensed intensional logic L_{*} , can be considered a pair $\langle\langle A, W, T, \leq \rangle, F\rangle$. F now assigns to constants their intensions, i.e. functions from points of reference $\langle w, t \rangle \in W \times T$ to their extensions at those points². Thus if α is a constant, its intension w.r.t. \mathcal{M} and g , $S(\alpha, g, \mathcal{M})$, is given by:

$$(8) \quad S(\alpha, g, \mathcal{M}) = F(\alpha).$$

A point of reference encodes the relevant aspects of a context of utterance - in Montague's theory these are, as far as this language is concerned, a world w and time t . The notion which is recursively defined is still the extension of a sentence (but now relative to a point of reference): $S(\alpha, g, \mathcal{M})(\langle w, t \rangle)$. Thus I shall also write, where φ is a sentence, $T_{*}(\varphi, g, w, t, \mathcal{M})$ for $S(\varphi, g, \mathcal{M})(\langle w, t \rangle) = 1$. A model in the old sense, i.e. an assignment of extensions to expressions, is thus a pair consisting of an interpretation \mathcal{M} and a point of reference $\langle w, t \rangle$.

The set of types is given an additional clause for intensions: iii) when $a \in Y$, $\langle s, a \rangle \in Y$. s is the type of points of reference. In L_{*} there are variables of every type, so if α is a variable of type a , $g(\alpha)$ will be a member of the set of appropriate extensions of type a , and

$$(9) S(\alpha, g, m)(\langle w, t \rangle) = g(\alpha).$$

For functional application, of a function β of type $\langle a, b \rangle$ to an argument δ of type a , we have:

$$(10) S(\beta(\delta), g, m)(\langle w, t \rangle) = S(\beta, g, m)(\langle w, t \rangle)(S(\delta, g, m)(\langle w, t \rangle)).$$

As well as the natural clauses for the various connectives we have a clause for \Box :

$$(11) \text{ If } \varphi \text{ is of type } t \text{ (then so is } \Box\varphi, \text{ and)}$$

$$T_*(\Box\varphi, g, w, t, m) \equiv (\forall w' \in W)(\forall t' \in T) T_*(\varphi, g, w', t', m)$$

(boxes are forever). The language also has an intension forming operator \wedge which forms an expression $\wedge\alpha$ of type $\langle s, a \rangle$, from an expression α of type a . $\wedge\alpha$ denotes the intension of α - thus if α is a sentence, $\wedge\alpha$ could be read as that α . So

$$(12) S(\wedge\alpha, g, m)(\langle w, t \rangle) = S(\alpha, g, m).$$

If α denotes an intension, ie. is of type $\langle s, a \rangle$, the result of applying the extension forming operator \vee to it denotes the corresponding extension, ie. $\vee\alpha$ is of type a , and

$$(13) S(\vee\alpha, g, m)(\langle w, t \rangle) = (S(\alpha, g, m)(\langle w, t \rangle))(\langle w, t \rangle).$$

There are a number of different ways one could characterise the syncategorematic expressions of L_* using IMT. In order of increasing strength: χ is syncategorematic if i) the expression is not explicitly assigned an entity or function, its sense; ii) there is no clause which (even) assigns χ an extension (showing its sense - assigning it an intension implicitly); iii) there is no clause which shows how the presence of χ affects the extension of a complex expression in which it occurs. Derivatively, these criteria apply to expressions δ of a natural language L which is interpreted by translation into L_* - ie. δ is syncategorematic if its translation $TR(\delta)$ is. (There is another, syntactic criterion for L : iv) δ is not a member of any set

B_A of basic expressions of category A, ie. if it is only introduced recursively in one of the syntactic rules for L.) Necessity, \Box , is categorematic by iii), but not by i) and ii). This is not essential - we can, eg. move \Box up a grade by using the clause for functional application 10) and the clause

$$(14) S(\Box, g, \mathcal{M})(\langle w, t \rangle) = \text{that function } h \text{ such that for any proposition } p, h(p(\langle w, t \rangle)) = 1 \text{ iff } (\forall w')(\forall t')p(\langle w', t' \rangle) = 1.$$

But clearly this is just a longwinded way of saying what 11) says. The point is the rather obvious one that if we are interested in stating the sense of an expression like \Box , it's no good simply assigning it some function as its sense, unless one also characterises that function in some other terms - and then talk of the function becomes a superfluous intermediary. Being syncategorematic à la i) or ii) therefore says less about the sense of a logical constant than the way it is being treated within the theory, viz., as a logical constant.

Central to Montague's conception of semantic theory is the claim announced in the title and first sentence of "English as a Formal Language": "I reject the contention that an important theoretical difference exists between formal and natural languages" ([73], p.188). The optimism in this claim is very similar to Davidson's faith in the capacity of Tarskian 'absolute' truth theory to be extended from formal to natural languages. Montague acknowledges this, but also that his claim is made from the perspective of a model theorist. If we take the paradigm of a formal language to be one whose semantics can be given in terms of FMT, we can be clear about the kind of difference between

formal and natural languages which Montague denied. His optimism, like Davidson's, heralds a programme: to demonstrate in detail that the techniques which enable the concept of truth to be defined for a formal language can be adapted to cope with the intensionality, indexicality, ambiguity, elipsis, vagueness, subjectivity etc. which infect natural languages.

Equally, if not more important than these specific differences between the two kinds of language is the simple fact that a formal language is defined into existence in front of our very eyes. In contrast to this, the workings of a natural language are not handed to us on a plate. This raises the question: what kind of evidence is a model theoretic treatment of a natural language accountable to? A lower bound on the acceptability of a structural semantic theory is its ability to cope finitely with all the different constructions in the language. Beyond that, the principle means of assessing model theoretic semantics is the theorist's reflective insight into the workings of his own language. One may use the word "intuition" here if that does not give the misleading impression that the process is not subjected to rational criticism. We have intuitions about scope distinctions, about the preferred scope of a word or phrase relative to another kind of phrase. These can count towards a systematic account of the kind of thing the former expression applies to, for example. We have intuitions about entailment relations, ie. about the relationship between the truth conditions of different pairs of sentences. Sometimes these intuitions are faulty, or people mistake their import. We shall see examples of this in section 2.3. We can discover mistakes because there are

principles, such as Grice's conversational principles, for the critical scrutiny of these intuitions - principles which enable us to skim off aspects of the significance of an assertion not strictly relevant to its conditions of truth. And of course, with certain qualifications, conjectures about intensions are subject to conformity with actual extensions. Some intuitions about meaning are (as) unsailable (as anything can be). All English speakers will agree that the word I refers, in a context, to the one who produces the sentence containing it in that context. Precisely how this simple insight is to be incorporated into a systematic formal theory is another matter. This will depend to a large extent on the best overall account of the ways contexts can regulate sense. The choice of such theories is subject to philosophical appraisal. As Dummett put it in his Shearman lectures, a semantic theory must serve as the basis for a theory of meaning - a model theory which does not present an acceptable picture of the workings of certain expressions has only technical significance. Possible worlds, for instance, provide a reasonable means of representing objective conditions of truth. But there are questions of detail to consider, as Kripke's work on rigid designation exemplifies (this leads back to intuitions about scope, of names in relation to modal operators - cp. §2.4). There are also questions of application - is this objectified notion really adequate for all the tasks we want sense to perform (which leads back eg. to the question of whether certain truth values assigned to propositional attitude sentences are acceptable)? Although Montague disdained psychological considerations, at least in so far as they translate into the precise requirement that

propositional attitude sentences be properly accounted for, they are a powerful constraint.

After semantic data regulated by philosophical inquiry, a second constraint on model theoretic structural semantics is syntactic data regulated by linguistic theory. Syntax and semantics exist in a certain degree of tension with one another. Syntax is a 'theory of meaningfulness' for a language - it should characterise all and only the expressions capable of bearing a meaning in the language. Traditionally, the logician demands of the syntax of a formal language that it a) defines the set of well-formed expressions and b) provides a basis for the semantic interpretation. Logical perfection in a language - as defined by Kaplan ([56], p.214) - requires that for each syntactic formation rule there is a corresponding semantic evaluation rule. An artificial language will be logically perfect unless there is some good reason not to design it so. But with a natural language syntax cannot simply be manipulated to serve the ends of semantics. First and foremost it is answerable to empirically given facts about well-formedness supplied by eg. native speakers' intuitions about grammaticality. Syntactic well-formedness and intelligibility or meaningfulness even seem to come apart at various places. Bearerless proper names and failed definite descriptions constitute one kind of example, where some philosophers have suspected that a well-formed sentence may fail to express a proposition. Another kind comes from the violation of conceptual possibility already alluded to above, and illustrated by Jones is a proof tree. But characteristic of such potential nonsense is its ability to produce sense when taken in an appropriate context - eg. in metaphorical

use, or embedded in a propositional attitude construction:

(15) Jones dreamt that he was a proof tree

(cp.[95],p.72). This contrasts with genuine syntactic rubbish, which is not capable of making sense - eg. in

(16)*Jones dreamt that he and not

I suggested earlier that structural semantics start from a basis in the major syntactic categories. It may then deviate from the syntactic classification either by singling out logical constants from their category or by subdividing these categories under pressure to account for certain inferences (there may be other pressures - cp. pp.109-110). The correspondence may also break down in the other direction, in that different syntactic categories are assigned the same semantic type (cp. p.99).

2: MODALITY AND MONTAGUE GRAMMAR

Montague's approach to English in PTQ was to interpret it indirectly, through a four stage process. i) The syntactic characterisation of a formal language L_* (and a system of logic within it), ii) The interpretation of L_* by a model theoretic truth definition, iii) The syntactic characterisation of a fragment of English and iv) The definition of a system of translation of sentences of that fragment to L_* , according to their syntactic composition. In this section I wish to concentrate on stages iii) and iv) of this process, with particular reference to the actual treatment of modality which Montague gave, and some extensions of it which allow other modal locutions to be treated. By supplying these extra stages, Montague demonstrated by example one way of making model theory applicable in the study of natural language.

The lexicon is divided into sets of different basic expressions, B_T , of basic terms, B_{TV} , of basic transitive verbs, and so on. T and TV are syntactic categories. The basic expressions of a category A form a subset of the phrases of that category; $B_A \subseteq P_A$. (There are no basic (ie. unstructured) sentences - B_t is empty.) There is a rule f which correlates every syntactic category with a semantic type, defined thus: $f(e) = e$, $f(t) = t$ and $f(A/B) = \langle\langle s, f(B) \rangle, f(A) \rangle$. Two syntactic categories may be correlated with the same type, as CN and IV, the categories of common noun and intransitive verb, are with $\langle\langle s, e \rangle, t \rangle$. Naturally, the categories CN/CN and IV/IV, of adjectives and verb-modifying adverbs, are¹ also mapped to the same type, $\langle\langle s, \langle\langle s, e \rangle, t \rangle \rangle, \langle\langle s, e \rangle, t \rangle \rangle$. A category A/B (and A//B) is

one of expressions which combine with an expression of category B to form one of category A. Thus necessarily, the only modal locution Montague deals with and sole occupant of $B_{t/t}$, the set of sentence-modifying adverbs, is correlated with the type $\langle\langle s, t \rangle, t\rangle$. This means that it is to be translated into an expression of L_* of this type - something which forms a sentence from a proposition-denoting expression. Where p is a propositional variable of L_* , type $\langle s, t \rangle$

(T1c) $TR(\text{necessarily}) = \lambda p(\Box^{\vee} p)$.

To be told that an expression is of category t/t is not yet to be told how it forms sentences out of sentences - we need a syntactic rule. Not surprisingly, it is this:

(S9) If $\delta \in P_{t/t}$, and $\beta \in P_t$, then $F_6(\delta, \beta) \in P_t$, where $F_6(\delta, \beta) = \delta\beta$.

F_6 is the 6-th structural operation, a function from expressions to expressions - in this case, concatenation. S9) has its corresponding translation rule

(T9) If $\delta \in P_{t/t}$, $\beta \in P_t$ and $TR(\delta) = \delta'$, $TR(\beta) = \beta'$, then

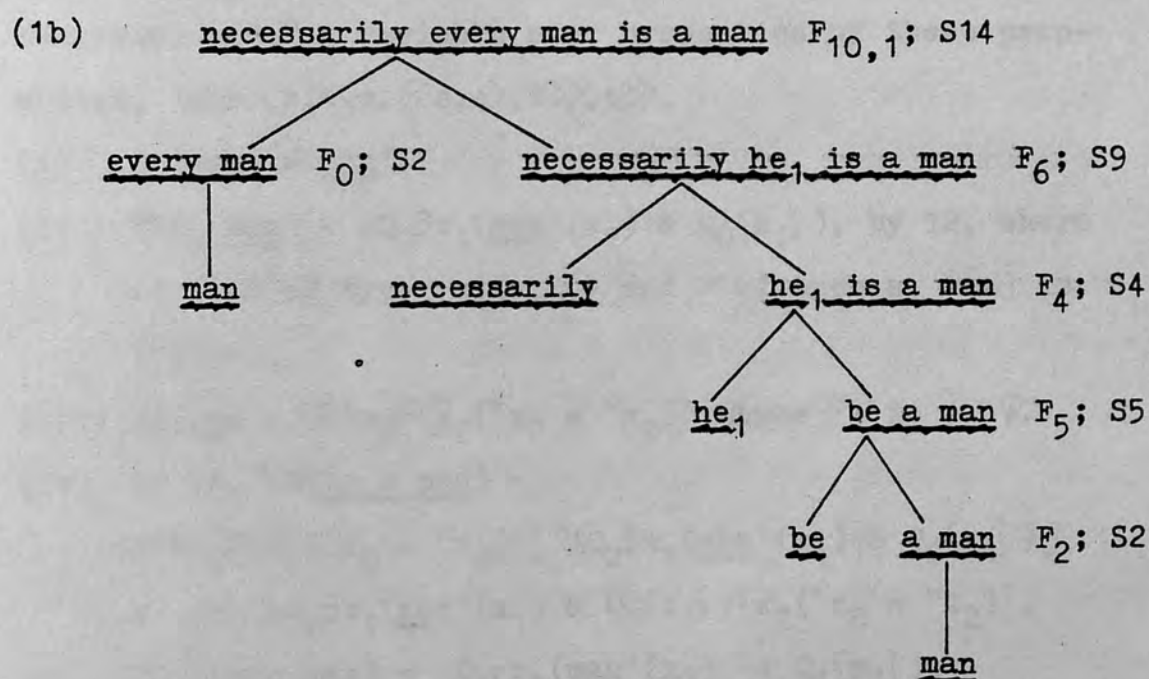
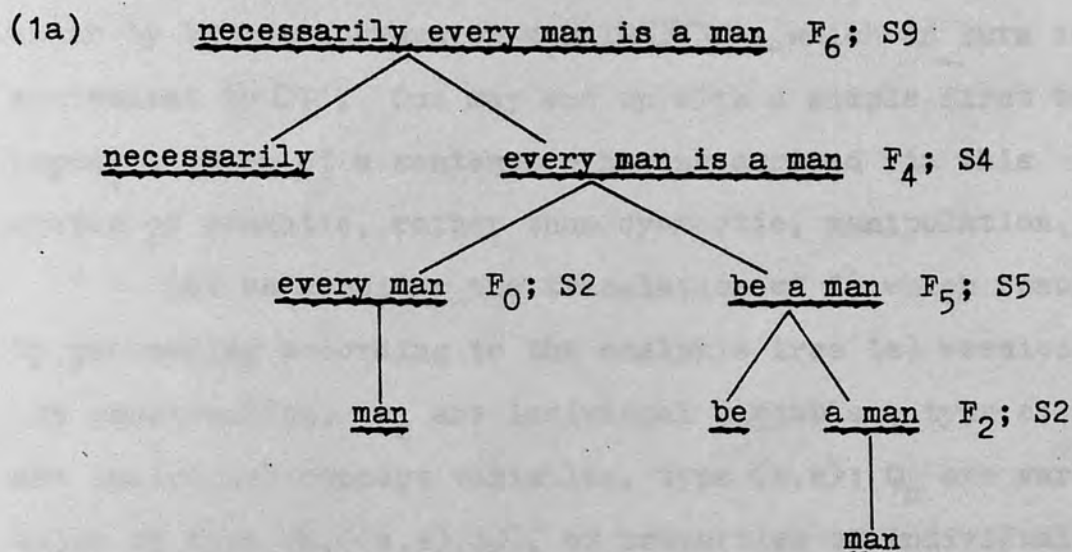
$$TR(F_6(\delta, \beta)) = \delta'[\wedge\beta'].$$

(To avoid unnecessary repetition, I shall adopt the convention here that φ' is to symbolise the translation into L_* of φ , and φ'' is to abbreviate "the result of replacing the main verb in φ by its first person singular present".)

Every sentence of the fragment will have as many analysis trees, which display its syntactic construction, as there ways of constructing it via the syntactic rules. Since there are non-trivial variants, this is to say that the fragment contains structurally ambiguous sentences. As an illustration, let us consider the de dicto and de re readings permitted of

(1) Necessarily every man is a man.

To the right of each non-basic expression displayed in the tree, I will list the structural operation and syntactic rule of PTQ by means of which it is produced from the expressions immediately below it. We have:



Translation proceeds according to construction, starting with the simplest and working up. It is at this point that one can appreciate the power and generality which the glut of variables in L_* (ie. from every type) imparts to the method of semantic representation through translation into L_* . The initial translations of non-basic English expressions generally involve a greater complexity than is

needed in their final representation, and are designed to be simplified by the application of a number of rules and meaning postulates. Thus if φ is a sentence of English,

$$\text{TR}(\text{necessarily } \varphi) = (\lambda p)(\Box \forall p)[\varphi'],$$

which by lambda conversion equals $\Box \varphi'$, which in turn is equivalent to $\Box \varphi$. One may end up with a simple first order representation of a sentence - having arrived via this system of semantic, rather than syntactic, manipulation.

Let us consider the translation of 1) which results by proceeding according to the analysis tree 1a) version of its construction. v_n are individual variables, type e; x_n are individual concept variables, type $\langle s, e \rangle$; Q_n are variables of type $\langle s, \langle \langle s, e \rangle, t \rangle \rangle$, of properties of individual concepts; and \mathcal{P} a variable over properties of these properties, type $\langle s, \langle \langle s, \langle \langle s, e \rangle, t \rangle \rangle, t \rangle \rangle$.

$$(i) \quad \text{TR}(\text{man}) = \text{man}'.$$

$$(ii) \quad \text{TR}(\text{a man}) = \lambda Q_0 \exists x_1 (\text{man}'(x_1) \& Q_0\{x_1\}), \text{ by T2, where if } \delta \text{ is of type } \langle s, \langle a, t \rangle \rangle \text{ and } \alpha \text{ of type } a, \delta\{\alpha\} \text{ is } (\forall \delta)[\alpha].$$

$$(iii) \quad \text{TR}(\text{be}) = \lambda \mathcal{P} \lambda x_0 \mathcal{P}\{\hat{x}_2 (\forall x_0 = \forall x_2)\} \text{ where } \hat{\alpha}\varphi \text{ is } \lambda \alpha \varphi.$$

$$(iv) \quad \text{By T5, TR}(\text{be a man}) =$$

$$\lambda \mathcal{P} \lambda x_0 \mathcal{P}\{\hat{x}_2 (\forall x_0 = \forall x_2)\} [\lambda Q_0 \exists x_1 (\text{man}'(x_1) \& Q_0\{x_1\})] \\ = \lambda x_0 \lambda Q_0 \exists x_1 (\text{man}'(x_1) \& Q_0\{x_1\}) \{\hat{x}_2 (\forall x_0 = \forall x_2)\}.$$

$$(v) \quad \text{TR}(\text{every man}) = \lambda Q_1 \forall x_3 (\text{man}'(x_3) \rightarrow Q_1\{x_3\}).$$

$$(vi) \quad \text{By T4, TR}(\text{every man is a man}) = \lambda Q_1 \forall x_3 (\text{man}'(x_3) \rightarrow \\ Q_1\{x_3\}) [\lambda x_0 \lambda Q_0 \exists x_1 (\text{man}'(x_1) \& Q_0\{x_1\}) \{\hat{x}_2 (\forall x_0 = \forall x_2)\}] \\ = \forall x_3 (\text{man}'(x_3) \rightarrow \lambda x_0 \lambda Q_0 \exists x_1 (\text{man}'(x_1) \& Q_0\{x_1\}) \\ \{\hat{x}_2 (\forall x_0 = \forall x_2)\} \{x_3\}). = \\ \forall x_3 (\text{man}'(x_3) \rightarrow \lambda Q_0 \exists x_1 (\text{man}'(x_1) \& Q_0\{x_1\}) \{\hat{x}_2 (\forall x_3 = \forall x_2)\}) \\ = \forall x_3 (\text{man}'(x_3) \rightarrow \exists x_1 (\text{man}'(x_1) \& (\forall x_3 = \forall x_1))).$$

$$(vii) \quad \text{So TR}(\text{necessarily every man is a man}) =$$

$$\Box \forall x_3 (\text{man}'(x_3) \rightarrow \exists x_1 (\text{man}'(x_1) \& (\forall x_3 = \forall x_1))).$$

man' is of type $\langle\langle s, e \rangle, t \rangle$, and man'* expresses the corresponding set of individuals, ie. is of type $\langle e, t \rangle$; man'* = $\lambda v \text{man}'(\wedge v)$. Since man is not an 'intensional' common noun, ie. it doesn't give rise to the Partee problem (cp. [73], p.248), this translation reduces, by Montague's second meaning postulate, to

$$\Box \forall v_3 (\text{man}'*(v_3) \rightarrow \exists v_1 (\text{man}'*(v_1) \& (v_3 = v_1))), \text{ ie.}$$

$$(2) \quad \Box \forall v_3 (\text{man}'*(v_3) \rightarrow \text{man}'*(v_3)),$$

a translation displaying the simple quantified modal logical form of 1), according to which it conveys the claim that the dictum that all men are men is necessary. Similarly, a translation of 1) in accordance with the analysis tree 1b) will simplify to

$$(3) \quad \forall v_3 (\text{man}'*(v_3) \rightarrow \Box \text{man}'*(v_3)).$$

Here is a generalised essentialist claim that something which is a man is so necessarily.

I want now to consider how to extend the Montague fragment in relation to the variety of forms of modal locution that occur in English. The question of how much natural modality can be interpreted in terms of the \Box which Montague defines at 2.1.11 above will be postponed until the next section. I will firstly consider three hypotheses for the treatment of the modal auxiliary verbs. i) As forming new intransitive verb phrases from old:

(S20) If $\delta \in P_{IV}$, then $F_{20}(\delta)$, $F_{21}(\delta)$, $F_{22}(\delta) \in P_{IV}$, where

$$F_{20}(\delta) = \text{must } \delta, F_{21}(\delta) = \text{can } \delta \text{ and } F_{22}(\delta) = \text{cannot } \delta.$$

Alternatively, one could expand the category $B_{IV//IV} = \{\text{wish to, try to}\}$ and treat them as normal IV-taking verb phrases. One way of securing the intended interpretation would be by laying down special meaning postulates, which

impose a restriction on those models considered admissible:

(4) $\Box(\text{must}'^{\wedge}\delta(x) \equiv \Box\delta(x))$, where δ translates any member of P_{IV} .

(5) $\Box(\text{can}'^{\wedge}\delta(x) \equiv \Diamond\delta(x))$, δ as in 4), etc.

This style of meaning postulate could also be employed in interpreting the adjective possible, ie.

(6) $\Box(\text{possible}'^{\wedge}\delta(x) \equiv \Diamond\delta(x))$, where δ translates any member of P_{CN} .

However, this is an unnecessarily roundabout way of achieving the same effect that would be secured by adopting special translations such as

(7) $TR(\text{can}) = TR(\text{possible}) = \lambda Q\lambda x\Diamond Q\{x\}$.

ii) A treatment which parallels the one Montague gives for tense and negation:

(S21) If $\alpha \in P_T$, $\delta \in P_{IV}$, then $F_{23}(\alpha, \delta)$, $F_{24}(\alpha, \delta)$, $F_{25}(\alpha, \delta) \in P_t$, where $F_{23}(\alpha, \delta) = \alpha \text{ must } \delta$, $F_{24}(\alpha, \delta) = \alpha \text{ can } \delta$ and $F_{25}(\alpha, \delta) = \alpha \text{ cannot } \delta$,

with the translation rule:

(T21) If $\alpha \in P_T$, $\delta \in P_{IV}$, then $TR(F_{23}(\alpha, \delta)) = \Box\alpha'(\wedge\delta')$,
 $TR(F_{24}(\alpha, \delta)) = \Diamond\alpha'(\wedge\delta')$ and $TR(F_{25}(\alpha, \delta)) = \Box\neg\alpha'(\wedge\delta')$.

iii) As a variant of the second approach, we have eg.:

(S22/T22) If $\varphi \in P_t$, then $F_{26}(\varphi) \in P_t$, where $F_{26}(\varphi)$ = the result of replacing the first verb δ in φ by must δ , and $TR(F_{26}(\varphi)) = \Box\varphi'$.

As will be apparent from 1b), adverbs classed as sentence-modifying are not precluded from taking scope narrower than that of a quantifier phrase. In PTQ distinctions of scope are held to be the responsibility of terms, eg. quantifier phrases, through the different ways they combine with expressions, rather than the responsibility of the operators with which they interact. This delegation of

responsibility is borne out by the following pair of sentences:

- (8) Any one of us may be killed
- (9) Every one of us may be killed.

Spoken before the troops are about to go over the top, 8) is naturally understood to assert of each of them that it is epistemically possible that he will be killed, whereas 9) makes the stronger claim that the following is possible: all of them are killed. Clearly, the different scope preferences cannot be a matter of the position of the modal term in 8) and 9), since it remains the same, but must arise through the ways the two quantifier phrases interact with it. Montague mentioned the fact that different quantifier words provide a means for reducing scope ambiguity in everyday language ([73], p.215), and suggested that rules might be given for restricting the number of possible syntactic analyses. Sometimes, of course, both scope readings may be possible.

- (10) Someone must be unhappy
- could be understood as a comment on the total misery in the world, with the epistemic necessity taking wide scope; alternatively, as a comment upon the screaming coming from the next room, with the existential quantifier out front.

These considerations must lead us to reject hypothesis i) concerning the modal auxiliaries. Treating an adverb as verb-modifying will not enable it to take scope wider than that of a quantifier phrase lying in subject position. Whereas it is acceptable to have possible modifying things of type $\langle\langle s, e \rangle, t \rangle$, as 9) and 10) testify, the corresponding restriction of must, can, may and co. would be incorrect. What of the converse policy of allowing

necessarily narrow scope? It may be thought, for example, that 1b) is an implausible reading of 1), given that every is a universal quantifier with a predilection for narrow scope. However, as the syntax of PTQ only allows a sentence modifying adverb δ to occur at the front of a sentence φ , we must allow for some means of syntactically transforming $\delta\varphi$ to insert δ into the middle of φ . We should reserve a de re reading for eg. "Every second is possibly your last" - for this reason as least, 1b) and its ilk should stay. (Note that since Montague adopts a once and for all domain of individuals A , the forms $\forall x\Box$ and $\Box\forall x$ are equivalent. Scope distinctions are significant because all the quantifier phrases in the PTQ fragment are restricted quantifiers. But because proper names are treated as rigid designators - through the meaning postulate $\exists u\Box(u = j)$, where j is the name in L_* for John - different syntactic analyses of such sentences as John must win have no semantic significance.)

Needless to say, we have sufficient means in PTQ to handle the distinctions of scope we saw arising in 16a) - 16c) of §1.4, as the following sample of (not fully executed) syntactic breakdowns displays:

- (a) $F_6(\text{possibly}, F_4(F_1(\text{man with the hat}), F_5(\text{be}, F_2(\text{candidate for presidency}))))).$
- (b) $F_{10,0}(F_1(\text{man with the hat}), F_6(\text{possibly}, F_4(\text{he}_0, F_5(\text{be}, F_2(\text{candidate for presidency}))))).$
- (c) $F_4(F_1(\text{man with the hat}), F_5(\text{be}, F_2(F_6(\text{possible}, \text{candidate for presidency}))))).$

In deciding between the two remaining proposals for the modal auxiliaries, it may help to bring in a few more constructions involving modality. There is a family of locutions which, like it is possibly true that φ , are based

on a dummy subject it. Here is a selection:

$$(11) \text{ It is } \left\{ \begin{array}{l} \text{possible} \\ \text{impossible} \\ \text{necessary} \end{array} \right\} \left\{ \begin{array}{l} \text{that} + \text{indicative} \\ \text{to} + \text{infinitive} \\ \text{for} + \text{infinitive} \end{array} \right.$$

$$(12) \text{ It } \left\{ \begin{array}{l} \text{may be} \\ \text{might be} \\ \text{must be} \end{array} \right\} \left\{ \begin{array}{l} \text{that} \\ \text{true that} \\ \text{the case that} \end{array} \right.$$

It is desirable that the relation is exhibited between these and other sentence-modifying adverbial phrases, such as John believes that. And there is another bunch of expressions, connected with the substantive possibility; for example

$$(13) \text{ There } \left\{ \begin{array}{l} \text{was} \\ \text{is} \\ \text{will be} \end{array} \right\} \left\{ \begin{array}{l} \text{genuine} \\ \text{real} \\ \text{slight} \end{array} \right\} \text{possibility} \left\{ \begin{array}{l} \text{that} + \text{indicative} \\ \text{of} + \text{gerund} \\ \text{of} + \text{noun phrase} \end{array} \right.$$

This apparently frank ontological commitment to possibilities is accompanied by a preparedness to count them: "one possibility is to ...", "there are three possibilities ...", etc.

Montague treated the verb phrases believe that, try to as unstructured - basic expressions of the categories $B_{IV/t}$ and $B_{IV//IV}$, respectively. Their grammatical role is given by functional application, eg. for the former

$$(S7/T7) \text{ If } \delta \in P_{IV/t} \text{ and } \varphi \in P_t, \text{ then } F_6(\delta, \varphi) = \delta\varphi \in P_{IV} \text{ and } \\ TR(F_6(\delta, \varphi)) = \delta'(\wedge\varphi').$$

An alternative would be to treat the that and to as syncategorematic, ie. with believe and try the members of $B_{IV/t}$ and $B_{IV//IV}$, and

$$(S7'/T7') \text{ If } \delta \in P_{IV/t} \text{ and } \varphi \in P_t, \text{ then } F_{27}(\delta, \varphi) = \delta \text{that} \varphi \\ \in P_{IV} \text{ and } TR(F_{27}(\delta, \varphi)) = \delta'(\wedge\varphi').$$

The it that φ construction of 11) and 12) can be produced

directly, via Bennett's rule (from [7], p. 145)

(S23) If $\delta \in P_{t/t}$, $\varphi \in P_t$, then $F_{28}(\delta, \varphi) = \text{it } \delta \text{ 'that } \varphi \in P_t$,

where in $B_{t/t}$ go such verbs as be possible.

(This forces adoption of the syntactically more superficial hypothesis iii) for the modal auxiliaries.) Montague's proposal has the disadvantage that it cannot handle

(14) John believes everything Mary believes.

This cannot be derived from the de dicto "John believes that everything Mary believes is true", for (14) could be true even though John knows nothing of Mary, and the latter could be true even if unwittingly, John disbelieves something believed by Mary. Nor is the other proposal any better just because it does not weld that to believe, for it still characterises believe as IV/t (type $\langle\langle s, t \rangle, \langle\langle s, e \rangle, t \rangle\rangle$), yet everything Mary believes is not a sentence. On this approach, one would have to postulate a second believe, member of B_{TV} , yet distinct from a third believe, as it occurs in John believes Mary (cp. for know). Moreover, in order to handle constructions such as (13), Montague would have to introduce a new category CN/t - for possibility that, belief that - ignoring the common role that plays in this and the previous construction (cp. to in be able to/ β and ability to/ β). Likewise, Bennett's rule passes over intuitively significant structure in be possible.) Consider, for example,

(15) There is a possibility that John will win.

With a rule for CN/t parallel to S7), we may expect

$TR(\text{possibility that John will win}) = \text{possibility that '}$

$(\wedge F_{win}(\wedge j))$.

With a rule for the there of existential assertions

(S24/T24) If $\alpha \in P_{IV}$ and α is of the form be δ , then $F_{28}(\alpha)$

$\in P_t$, where $F_{28}(\alpha) = \text{there } \alpha'$ and $TR(F_{28}(\alpha)) = \lambda Q \exists x Q\{x\}(\alpha')$ and Montague's relation notation $\delta(\beta)(\alpha) = \delta(\alpha, \beta)$,

$$TR(15) = \exists x \text{ possibility that'}(x, \wedge F \text{ win'}(\wedge j)).$$

We could introduce a special meaning postulate to reduce this to

$$(16) \Diamond F \text{ win'}(\wedge j),$$

but this would be ad hoc, and it would leave unexplained the relation between other occurrences of the noun phrase, as in

(17) The possibility that John will win amuses Bill and the concept of possibility, ie \Diamond .

It seems to me that in the respects we have just been discussing, the syntax of PTQ should be regarded as a simplification of the proper theory, one which exploits more fully the rich semantic resources available. The full theory treats as term phrases both that ϕ and everything Mary believes, but terms at a new level of ontology, the propositional level. The general idea, suggested by the work of E.B. Delacruz ([28]), is to make more use in the syntax of the multiplicity of semantic types by introducing a new level of grammar wherever terms are discerned which denote things of an appropriate type. In PTQ, everything is geared to the individual concept level, in that term phrases and common nouns respectively denote individual concepts and sets thereof. Delacruz adds to this the propositional level, which contains terms such as that John will win and common nouns such as proposition itself. Generalising from this, we may think of the syntactic rules of PTQ as, on the whole, formulating general modes of syntactic combination of eg. T's with IV's, no matter what the level of the expressions involved. Each syntactic category A is potentially

subdivisible into many A_n - let T_1 be the individual concept level terms, T_2 be propositional level terms, and T_3 be property level terms (eg. the capacity to win). The syntactic category of entity divides into e_1, e_2, e_3, \dots , and we modify the definition of the function correlating syntactic categories and semantic types as follows:

$$f(e_1) = e; f(e_2) = t; f(e_3) = \langle\langle s, e \rangle, t \rangle; \dots$$

Just considering the first two levels, we see that the category of TV now splits four ways: B_{IV_1/T_1} , the old set $B_{IV/T}$; $B_{IV_2/T_2} = \{\text{be}, \text{entail}\}$; $B_{IV_1/T_2} = \{\text{know}, \text{believe}, \text{ignore}, \dots\}$; and $B_{IV_2/T_1} = \{\text{amuse}, \text{surprise}, \dots\}$, as in 17).

The translation of most expressions will depend on their precise level. This is particularly evident in the case of those key words and constructions whose particular meaning is given by translation into the logical vocabulary of L_* and which may recur at different levels. For example, we might represent the translation of the existential there as follows:

$$(T24') \text{ TR}(F_{28}(\alpha), 1) = \lambda Q \exists x Q\{x\}(\wedge \alpha'),$$

$$\text{TR}(F_{28}(\alpha), 2) = \lambda J \exists p J\{p\}(\wedge \alpha'),$$

$$\text{TR}(F_{28}(\alpha), 3) = \lambda \mathcal{P} \exists Q \mathcal{P}\{Q\}(\wedge \alpha'),$$

where J is a variable over properties of propositions, type $\langle s, \langle\langle s, t \rangle, t \rangle \rangle$. Note that since this restricts there to combination with the verb be, to recover such phrases as there may be, hypothesis iii) concerning the modal auxiliaries is forced upon us. Consider the way apposition functions at different levels. As well as Montague's treatment at the individual concept level ($S3/T3$), we have

$$(S25/T25) \text{ If } \beta \in P_{CN_2} \text{ and } \varphi \in P_t, F_{29}(\beta, \varphi) \in P_{CN_2}, F_{29}(\beta, \varphi) =$$

$$\beta \text{ that } \varphi \text{ and } \text{TR}(F_{29}(\beta, \varphi)) = \lambda p (\beta'(p) \ \& \ p = \wedge \varphi')$$

$$\text{If } \beta \in P_{CN_3} \text{ and } \delta \in P_{IV}, F_{30}(\beta, \delta) \in P_{CN_3}, F_{30}(\beta, \delta) =$$

β to δ and $TR(F_{30}(\beta, \delta)) = \lambda Q(\beta'(Q) \ \& \ Q = \wedge \delta')$.

Examples of B_{CN_2} are belief, proposition, possibility and truth; examples of B_{CN_3} , ability, tendency.

With the addition of a few natural special translations we are able to reduce a variety of English modal sentences to the same underlying form. For example, the following is a plausible rendering of the substantive²

(T1f) $TR(\text{possibility}) = \lambda p(\Diamond \vee p)$

- just the translation one would expect that Montague would have given for possibly. But how one assembles an assertion with these two words differs. Using T25), T1f) and the propositional level translation for the indefinite article

$TR(\text{a possibility that John will win}) =$

$\lambda J \exists p(\Diamond \vee p \ \& \ p = \wedge F(\text{win}'(\wedge j)) \ \& \ J\{p\}).$

Using T24') and the propositional level translation of be (cp. [28], p. 188)

(18) $TR(15) = \exists p(\Diamond \vee p \ \& \ p = \wedge F(\text{win}'(\wedge j)))$

which reduces to 16), or $\Diamond F(\text{win}'_*(j))$ - something which could have been said in a less long-winded manner by possibly John will win.

For it is possible that φ and its variants we have to appreciate the workings of adjectives in predicative position. To classify adjectives as CN/CN is to take their attributive use, α is a $\delta \beta$, $\delta \in P_{CN/CN}$, $\beta \in P_{CN}$, as fundamental. This must be reconciled with the fact that they are used predicatively, α is δ , since the semantic type(s) correlated with CN/CN deems them incapable of applying directly to the relevant kind of individual. What makes the attributive construal plausible as a more general account than taking the predicative use as basic is the fact that certain adjectives, eg. former, only seem to make sense used

attributively. One may then take α is δ as derivative from α is a $\delta\beta$, with β a dummy noun appropriate to the level: entity, proposition, etc. We can secure the intended interpretation of the predicative possible through a meaning postulate:

$$(19) \Box(\text{possible}'(\wedge \text{proposition}')(p) \equiv \Diamond p).$$

(This cannot be generalised to all β ' translating some member of P_{CN_2} : a possible belief is a possible-to-be-enter-tained belief, rather than a possibly-true belief.) One way to put this into action is to use a deletion rule:

$$(S26/T26) \text{ If } \beta \in P_{CN_2/CN_2}, F_{31}(F_5(\text{be}, F_2(F_6(\beta, \text{proposition})))) \\ = \text{be}\beta \in P_{IV_2}, \text{ and } TR(F_{31}(\varphi)) = TR(\varphi).$$

Using Delacruz's rules for that-clauses,

$$(S27/T27) \text{ If } \varphi \in P_t, \text{ then } F_{32}(\varphi) = \text{that}\varphi \in P_{T_2}, \text{ and} \\ TR(F_{32}(\varphi)) = \wedge J J\{\varphi'\},$$

we can construct that John will win is possible using S4).

This too will simplify to the RHS of 18), and so to $\Diamond F(\text{win}' * (j))$. Finally, it is plausible to take the it δ that φ construction as derived from this form of sentence by reversing the order of subject and verb and slapping a dummy subject on the front:³

$$(S28/T28) \text{ If } \alpha \in P_{T_2} \text{ and } \alpha = F_{32}(\varphi) \text{ for some } \varphi \in P_t, \text{ and } \delta \in \\ P_{IV_2}, \text{ then } F_{33}(F_4(\alpha, \delta)) = \text{it}\delta'\alpha \in P_t, \text{ and for any} \\ \beta TR(F_{33}(\beta)) = TR(\beta).$$

(This construal of the dummy subject it does not force a choice between hypotheses ii) and iii) above.)

3: MODALITY: MEANING

I now turn to what, from the reflective native's point of view, some of the various modal expressions mean. We can then evaluate the claim of systems other than the S5 of Montague's intensional logic to represent them. One motivation for employing S5 is for the benefit of theory: it is needed for the meaning postulates to be stated with full rigour. It is a separate issue whether we should translate ordinary modal locutions into the operators of that system. Translation in terms of S5 has some plausibility in the case of the particular word which Montague treated. One might argue that because we have few intuitions about modal reduction principles such as necessarily φ \rightarrow necessarily necessarily φ , we need a system much weaker than S5¹. But this way of thinking ignores what would be the more basic fact, the failure of intuition arising through its lack of a grip on the combinations such as necessarily necessarily φ which the reduction laws employ. If we ordinarily find no point in such iterations, having no use for them, our everyday necessarily should be represented by a system which collapses this kind of modal distinction, through reduction laws like those of S5, rather than one which lets them multiply².

It may nevertheless be that a more complex account of necessarily will be required than the one Montague gives. After all, in everyday life we have little use for the logical/metaphysical/conceptual notion of necessity. It is often used to express a strictness of conditionality of a broadly causal nature, as for example in "Reflation will not necessarily lead to inflation". However, I would prefer us

to turn our attention to the meaning of the modal auxiliaries can and must, as these provide, in a clear way, a rather rich variation on the theme of modality, and raise interesting questions about the variation of meaning from context to context. They have recently been studied in this connection by Angelika Kratzer. She poses the question: how many different senses of must and can should we recognise? One might first distinguish a small number of major categories - the ones I shall shortly discuss are A) Epistemic, B) Existential/Universal, C) Ability/Opportunity and D) Deontic. One might then think that each of these categories further subdivides. On this point appeal might be made to the authority of Aristotle: "the senses in which we should call a boy and a man and a eunuch 'incapable of begetting' are distinct" (from the *Metaphysics*, $\Delta 12, 1019b$). Kratzer rejects this idea that can and must are multiply ambiguous, favouring the hypothesis that they are univocal but relative in sense. But before considering her proposal, it will be helpful to review some of these major categories of their sense.

A) Epistemic. It is widely recognised that epistemic possibility and necessity are frequently involved in everyday modal discourse. Quine, for example, writes in *Word and Object* (p.195) that possibly is usually "a modestly impersonal rewording of ... 'I am not sure but what'." Of the modal auxiliaries must is an obvious representative, as in Karttunen's example ([57])

(1) John must have left.

By inserting must the speaker draws attention to the epistemic status of the proposition in question -

(2) John has left

- that 2) follows as an inference from the information he possesses, ie. is not known directly. The epistemic use of can is more or less restricted to the negative (cp.[75], § 8.4), may being the favoured idiom, but it seems to me that an example Palmer gives as neutral can be understood as epistemic:

(3) Who knows? It can go either way.

Imagine it used of a close run ballot, when all the votes are in (so that the result is determined), yet not fully counted, to convey the epistemic uncertainty of its utterer.

Karttunen formulates a rule, a "constraint on the use of possible that we all seem to follow ...":

(4) Whatever is cannot (possibly) be otherwise.

To motivate the principle, we are presented with some examples, not syntactically ill-formed, but anomalous as "a matter of the meaning of possible" ([57], p.5). Here are two:

(5) *It isn't raining in Chicago, but it may be raining there

(6) *John is mistaken, but it is possible that he is right.

It is not entirely clear that the second has no acceptable meaning - one might say "It is possible that John is right", as a concession, "but (in fact) he is mistaken". This is a quibble, since it is to find 6) a reading in terms of logical possibility, and this is not the concern of principle 4). What is more serious is the unclarity of the status of the principle, of the level of explanation offered.

Karttunen considers the consequences of adding 4) to one of the standard accounts of the truth conditions of $\Diamond\varphi$, where the possibility is logical - not surprisingly, it leads to a collapse of modal distinctions - by way of rejecting such an account for may and possible. This suggests that the explanation is offered at the level of sense³. For

other phenomena, however, he clearly gives an explanation in terms of conversational principles. An example of this is the Aristotelean distinction between one-sided and two-sided possibility. If \Diamond_2 represents the latter, then $\Diamond_2\varphi \equiv \Diamond_2\neg\varphi$. For one-sided possibility, $\Box_1\varphi \rightarrow \Diamond_1\varphi$. Aristotle recognised that these two notions cannot be combined, for then $\Box\varphi \rightarrow \Diamond\neg\varphi$ would ensue. Karttunen argues that if someone says something of the form $\Diamond\varphi$ and is obeying the Gricean conversational maxim 'Make your contribution as informative as is required', he indicates that he doesn't know the facts for sure. So for all he knows, $\neg\varphi$ is also possible.

This seems a good explanation to me, and it illustrates one of the great dangers in using intuition when dealing with matters of sense. The fact that we don't or shouldn't assert $\Diamond\varphi$ when we know φ or even $\Box\varphi$, so that we only assert it on occasions when $\Diamond\neg\varphi$ is also assertible (or at least cannot be gainsaid), does not mean that this is to be laid down as a principle, ala \Diamond_2 , governing our concept of possibility. Exactly parallel considerations apply in the case of the principle 4). By Grice's maxims of Quality, 'Do not say what you believe to be false', 'Do not say that for which you lack reasonable evidence', a speaker should have good grounds for an assertion φ he makes. Even if he's not absolutely certain that φ , the epistemic possibility of $\neg\varphi$ must be a remote one. So being in a position to assert φ makes maybe $\neg\varphi$ conversationally inappropriate. But this is not to say that if φ is true, whether we know/can assert it or not, it is epistemically impossible that $\neg\varphi$. As stated, principle 4) makes English speakers out to be either epistemologically very arrogant, or extreme idealists. But since Karttunen also comes to the conclusion that examples

such as 5) and 6) are not actually inconsistent but violate "the rules of discourse", one cannot definitely accuse him of adhering to it in this spirit. It seems that he is unaware of the conflict between the two levels of explanation offered.

Let us now consider the epistemic must. Karttunen seems to concur with the "intuitive feeling that [1)] is a weaker assertion than [2)]" (p.13). A distinction perhaps worth drawing here is between the factive and the committal. Call a sentence ψ factive with respect to φ if the truth of ψ requires, inter alia, the truth of φ (there are other, pragmatic, definitions of factivity, but these may be ignored for our purposes.) Call a sentence ψ committal w.r.t. φ if the sincere assertion of ψ commits the speaker to the truth of φ . The former is a subclass of the latter, and φ is committal, because factive, w.r.t. φ . It seems clear that 1) is at least committal w.r.t. 2) - we use the must form as a way of asserting the unmodalised sentence. It is rather like saying "I conclude that John has left". But there is a difference. Someone might truly assert the latter on the basis of some erroneous and idiosyncratic theory of John's behaviour, but we would regard their assertion of 1) as false if 2) fails to follow from generally acceptable assumptions - even if their idiosyncratic conclusion turns out to be correct. The must idiom is not as subjective as personal reports of belief. But nor does it appear to be as strong, as objective, as knowledge. If someone makes a must judgement on the basis of an inference that would be generally accepted as practically certain, but is unfortunate enough to hit a rare counterexample, we would not readily call that judgement false. The must idiom is

not factive - it is a guarded form to be used when the speaker does not know for sure.

Since our information changes through time - we gain and lose knowledge - we should expect time to enter crucially into the definition of the notion of epistemic necessity [e] by means of which we interpret must, may, etc. The importance of this feature is brought out by Strawson's recent analysis of might have statements ([98]). In one sense, might have would naturally be translated in terms of the combination $\Diamond P$, as in "It might have rained last Sunday (I can't remember)". But Strawson is interested in another sense of the phrase, one which is compatible with our present certainty that the event in question did not happen. A good example he gives of this is "You might have been killed" - spoken, say, to someone who has just avoided a fatal accident. We can approximate his analysis within the Montaguean style as follows: $TR(\alpha \text{ might have } \delta ''') = P \Diamond F \alpha'(\wedge \delta')$ (where $\delta \in P_{IV}$ and δ''' is its past participle form). A might have statement is true in this sense if there was a time when it was not certain that α wouldn't δ . But certain to whom? Strawson argues that the knowledge concerned would be of an ideal information gatherer, one who would collect facts available at the past point in time, even if they were not then collected, and who is equipped with our present knowledge of the way the world develops. $\alpha \text{ might have } \delta'''$ will be true if we can imaginatively project ourselves back, in the guise of this spy, to a point in time when the particular facts available at that time would not lead us to say, on the basis of our current knowledge of general truths, that the event or state in question will not occur. Obviously we do not import particular knowledge

concerning the fact of α 's (not) δ -ing. In making these might have judgements, we are not interested in the possibly erroneous perception of how things would develop of people around then, if indeed there were any at all. Perhaps no one was then appraised of all the relevant particular facts; perhaps no one then knew the laws we know about what ensues from what. Strawson does not consider the import of past facts which would have a bearing on the matter, then known but now lost. This perhaps reflects the fact that these cannot enter into our (current) appraisal of the situation. We must assume that our spy would not be empowered to collect them.

We can summarise these findings by providing a suitable model theoretic characterisation of \boxplus :

$$(7) T_*(\boxplus\varphi, w, t, \mathcal{M}) \equiv (\forall w') w \underset{t}{\approx} w' \rightarrow T_*(\varphi, w', t, \mathcal{M}).$$

Worlds are now to be considered as splayed out in time - they are possible world histories. $w \underset{t}{\approx} w'$ is in effect the alternativeness relation relativised to time: w' is an epistemic alternative to w at t . Since our knowledge changes in time, from $w \underset{t}{\approx} w'$ we can infer nothing about $w \underset{t'}{\approx} w'$ for times t' earlier or later than t - there is branching backwards and forwards in time. We also established that neither $\varphi \rightarrow \boxplus\varphi$ nor $\boxplus\varphi \rightarrow \varphi$ is a principle which governs the concept of necessity underlying the epistemic must. So $\underset{t}{\approx}$ is not reflexive, though it would appear to be symmetric and transitive. As Palmer has observed ([75], p.19), combinations of epistemic modalities, on those infrequent occasions when they naturally occur, do intuitively seem to be pleonastic. With, for example

(8) It may possibly rain tomorrow
one feels that the extra modality is not there to mark a

finer distinction, but is essentially redundant. Likewise, you would not say "It may be that John must have left" if you're not convinced that "John has left" is correctly inferrable from the available information; you would simply deny that "John must have left" is true.

B) Existential/Universal. Palmer has an example

(9) The squid of the genus *Loligo* can be as much as two feet long,

which, along with Ryle's

(10) Stones can float (because pumice-stone floats)

suggests that in combination with a generic or a bare plural, can may be used as an existential quantifier. 9) "does not mean that a squid may vary in length from time to time, but that some squids are two feet long"([75],p.153). As regards must as a universal quantifier, he gives

(11) All scientific results must depend on a rather specialised form of history,

writing that "virtually the same meaning" would result if either the all or the must were dropped.

It seems to me that the examples 9)-11) do not exemplify a distinct category of sense or usage for the modal auxiliaries. The phenomenon is principally connected with the generic and bare plural forms. For one thing, 10) must also have a universal reading, since it has the same form as

(12) People over 18 can vote,

which is naturally understood as: All people over 18 can vote, in which the modality is not redundant. So although the bare plural form can be read either as existential or universal, the insertion of can does not force the former reading. Furthermore, we can find examples of the generic

construction, eg. "The NASA space shuttle can re-enter the earth's atmosphere intact" which would have been true if the things had been built but never tested, even though the corresponding existential paraphrase (some ... do/will) would be false. The element of potentiality expressed here by the can cannot be captured in terms of \exists . Similarly, although Palmer writes that a generalisation with must is roughly synonymous to one without, in effect he concedes that the former will be stronger than the latter⁴.

C) Ability/Opportunity. Under this category falls Kratzer's 'dispositional' must: If you must sneeze, at least use your handkerchief. Its suitability for inclusion with the can's of this category, on which I shall concentrate, is brought out by the paraphrase she gives of it: "If you cannot help sneezing,...". Much has been written on the can of ability, and I shall merely aim to highlight some of its temporal features. I think there is considerable intuitive support for the idea that there are basically two related can's falling within this category, those introduced by Vendler in the following passage.

There are people who can drink a gallon of wine in one draught. Suppose one of them has performed that remarkable feat a minute ago. Then it is quite unlikely that he can do it again now. Should we say then, at this moment, that he can, or rather that he cannot, drink a gallon of wine in one draught? He can and he cannot.
([105], p.116)

The sense in which he can, which Vendler labels can₂, seems to correspond to what Honoré ([52]) distinguishes as the general can, which relates to types of actions. The sense in which he cannot, can₁, is a particular can relating to a particular occasion. It is natural to construe this distinction in terms of tense. can₂ should be understood as the simple present, since it relates to the possession, over

a period of time which includes the present, of a general ability or freedom from impediment to perform some kind of action or be in some kind of state. Exemplifying action: "He can drink a gallon of wine in one draught", exemplifying state: "You can believe in disarmament without being a unilateralist". can₁ is naturally construed as future, since it relates to the ability to perform, occur, etc. on some particular occasion, as in Austin's example "He can sink this putt". An example of a state: "We can stay at home this evening". Note that if he fails to sink it, it doesn't necessarily mean that that prediction was false, since an ability, whether general or in relation to some specific action, can tolerate a failure or two (cp.[3],p.218, [52], p.466). The distinction can₁/can₂ is not always clearly marked in ordinary language, since the same form of words can do duty for either sense. "If I work hard, I can get away by five" may refer to the speaker's general situation at work, or may be used with reference to his particular predicament on that day. There is a difference: if true in the general sense, then true with respect to that day, though not conversely. As Vendler's example shows, the inference from can₂ to can₁ cannot be generally valid, in that the performance of a feat may temporarily incapacitate the agent. (Honoré also distinguishes a third quite particular can, one which I have difficulty in recognising. It is virtually equivalent to will, and predictions employing it are therefore supposed to stand or fall with the success or failure of the particular action in question - ie. it is intolerant of failure. It seems to me that when someone holds "He can do it" false when the person in question fails to produce the goods, they are construing the failure as

evidence for the lack of the ability, in one of the two senses already distinguished.)

If this sketch is along the right lines, a full treatment of can will only arise within the context of a systematic treatment of aspect. Nevertheless we can bring out certain salient features of the word without a proper theory of aspect. For example, the importance of the temporal parameter in its semantic evaluation is well illustrated by an argument devised by Lehrer and Taylor. They use this argument to cast doubt on the principle P that "If A is a certain means to some end and an agent can do A, it follows that the agent can achieve that end" ([114], p.399). More abstractly, a question mark against P would mean that the rule $\Diamond\varphi, \Box(\varphi \rightarrow \psi) \vdash \Diamond\psi$ would be in jeopardy. Their argument concerns a person Smith who, though nothing prevents him catching a plane at 3.30, chooses not to do so. The plane is the only means of getting to a certain city by four o'clock. The argument runs as follows:

- (13) If Smith does not leave at 3.30, then he cannot arrive at 4.00 ,
- (14) Smith does not leave at 3.30 , therefore:
- (15) Smith cannot arrive at 4.00.
- (16) If Smith leaves at 3.30, he will arrive at 4.00 ,
- (17) Smith can leave at 3.30 , so by principle P:
- (18) Smith can arrive at 4.00.

In 15) and 18) we appear to have reached a contradiction.

Faced with this result, we may either challenge one or more of the premises, query whether 15) and 18) are genuine contradictories, or reject the principle P. The second response has its appeal - as in Vendler's example, we might say "There's a sense in which he can arrive, and a sense in

which he can't". But this is not obviously a case of can₁/can₂ ambiguity, and simply to level the charge of a fallacy of equivocation is not yet to explain anything. The first response will provide for us a route to such an explanation, for we may ask whether the premises add up to a description of the situation consistent with any one particular temporal perspective on it. In fact the first thing we should ascertain about an argument like this is the use of tense it involves. The most plausible construal is that it relies on that use of the simple present with which one may narrate a past sequence of events. Typically, the ordering of sentences in the narrative will follow the temporal ordering of the events. In this respect the second inference 16)-18) should come first. It takes us back to a point before 3.30. The first inference moves us on past 3.30. Premise 14) is the most concrete of the lot, reporting Smith's non-departure at 3.30. Now, as Lehrer and Taylor recognise, what we are able to do changes with time - in particular, one can lose an ability. Up to 3.30, Smith has the ability to arrive at 4.00, since it is assumed that it is a free choice he makes at 3.30, ie. not determined by something occurring prior to that time. But he loses the ability once 3.30 is past and he is not on the plane. So the 'senses' in which the conclusions 15) and 18) differ are basically the same as in Kant's example "He is young" and "He is not young" said of the same person at different times (Critique of Pure Reason, B192). Before 3.30, "Smith cannot arrive at 4.00" is false - we cannot then detach that as consequent of 13), because we cannot then assume that Smith will not leave at 3.30, even if that turns out to be the case. But after 3.30 it becomes true and 18) becomes false. "Smith could arrive

at 4.00" will still be true, but that does not contradict it.

We can capture the temporal relativity of can which this example brings out in terms similar to the clause that would parallel 7) for \Diamond . But more conditions will be placed on \approx for the notion \Diamond of historical possibility which it defines⁵. \approx_t is to be an equivalence relation, and if $w \approx_t w'$ then for any t' earlier than t , $w \approx_{t'} w'$. $w \approx_t w'$ represents coincidence of worlds w and w' up to t . Two such worlds may diverge at a subsequent time, and consequently $\Diamond F\varphi \rightarrow F\Diamond\varphi$, though not its converse, is invalid in this system. This captures Smith's predicament accurately enough, since he loses an opportunity. It may be that \Diamond basically suffices for can₁, for here the notion of ability expressed coincides roughly with the existence of some relative possibility. If a tutor says "I can see you at six o'clock tomorrow", and then checks in his diary only to find he is engaged, he might then say "No I can't see you then, I'm seeing someone else", the inability is an impossibility, not absolute but highly relative to the fulfillment of previous commitments and so on. Often we use can in this way when the action is not really impossible but highly inconvenient or awkward - the implication being that there is not really any choice in the matter, that there is not really any other option ("I can't" rather than "I won't").

Intuitively, can₂, when used in connection with actions, has far less to do with the existence of a possibility, however relative, than can₁. For one thing, the mere existence of a possibility does not seem sufficient for the truth of a can₂ statement. We don't hold "He can swim" to be true of someone at the water's edge who has not yet

learnt to swim but is about to do so. We gain abilities in the sense of can₂, not just lose them. Moreover, if someone actually but fortuitously executes some difficult action, or performs part of a complex procedure, we would be reluctant to say that he can do the thing in question. If someone manages to repair a TV by poking around in the back of it and knocking a loose component into place, we wouldn't say "He can repair TV's" or even "He can repair this TV", if he lacks the requisite training. What counts with kinds of actions is the possibility of repeatedly performing them on demand - with allowance for the odd failure. This is basically the idea of habitual performance which the simple present expresses and which is perhaps impossible to accurately represent formally. "He repairs TV's" is true now if there is a reasonable period of time which includes the present and in which there are times at which a corresponding progressive form can be truly asserted: "He is repairing a TV", and during which moreover some kind of conditional is true, roughly, that when he works he repairs TV's. The repairings are not totally fortuitous. "He can repair TV's" relates to the possibility of: he repairs TV's, ie. to the possibility of habitual repairings. (It is for this reason that an analysis of can₂ in terms of a subjunctive conditional is so plausible - and more manageable. We shall consider conditionals in connection with dispositions, in Chapter 3.) Normally, our grounds for asserting $\propto \text{can}_2 \delta$ are that the agent has δ -ed, but this is not essential. We can imagine "He can repair TV's" truly asserted of the apprentice technician who has so far learnt only theory and has not yet put his ability to use. It should be noted that this kind of analysis is inappropriate for stative verbs.

Statives lack the progressive form (cp.[105],pp.102-3). The truth of "John loves Mary" over a period of time requires that it is true at any instant in that period. "John can love someone (if they love him)" relates not to repeatable occurrences but to the possibility of his being in a certain state.

D) Deontic. Kratzer gives as an example

- (19) All Maori children must learn the names of their ancestors.

It seems we could also include here what she calls the 'preferential' must:

- (20) When Kahukura-nui died, the people of Kahungunu said:
Rakaipaka must be our chief.

It expresses in part the fact that Rakaipaka is the only person they will permit to be chief, just as "Rakaipaka can be our chief" could be used to express a permission, a deontic can, granted in virtue of their preferences. Alternatively, it could be understood as a can₁. It has often been noted that the borderline between this can and the deontic is not at all clear cut - Palmer cites the reporting of rules and regulations as constutable either way ([75], p. 149). When, for example, I say what must can express, I talk about a possible use within English, ie. what is permitted by the rules of the language. It has been suggested, with some plausibility, that our conception of necessity and possibility stems from deontic notions of obligation and permission. The one major difference between this category and the others is that we may be empowered to grant permission or lay down an obligation - the deontic can and must have a performative use. Assertoric use of must and can can intuitively seem to report obligations or permissions

granted from a non-human authority. must may express a conceptual obligation, eg. in "We must therefore conclude," or "If you accept an intuitionist theory of meaning, you must reject the principle of bivalence". The deontic must shares a feature we decided was possessed by its epistemic counterpart, namely, that it doesn't necessarily follow that what must be is. Deontic modality, which must be tensed to allow for change in obligation over time, is characterised by the axiom $\Box\varphi \rightarrow \Diamond\varphi$. Deontic alternativeness is serial: $(\forall t)(\forall w)(\exists w')w \underset{t}{\approx} w'$ (cp. [11], chapter 6).

We can now return to the question of the overall sense of can and must. Kratzer musters two lines of thought against the view that they are ambiguous. The first depends on the assumption by the ambiguity view of the 'Aristotelean' move to distinguish different senses of the verbs within each major category. Since our dispositions, epistemic states and so on are constantly changing, the ambiguity view would need a never-ending list of lexical entries must₁, must₂, must₃, "But we might not have enough numbers," she reasons, "And even if we had enough numbers, it would not be very sensible to use them here. In everyday conversation we do not use subscripts when we use the words must and can. Somehow we do without them." ([58], p.340) The second line of thought is that in addition to this multiplicity ("billions") of must's, the ambiguity theorist would have to recognise a neutral must, to be found in such sentences as

(21) In view of what their tribal duties are, the Maori children must learn the names of their ancestors

(22) In view of what is known, the ancestors of the Maoris must have arrived from Tahiti.

Her proposal is that when the relativity is unpacked through one of these in view of relative clauses, the must (or can) we are left with is neutral, and expresses the kernel of meaning which is the common element running through each occurrence of the verb - as, for example, in 21) and 22) minus the qualifying phrases (cp. 19)). The in view of phrase spells out explicitly what would otherwise be provided implicitly through context, giving rise to the impression of difference in sense. The phenomenon is not ambiguity but a relational sense of must and can, requiring completion in one of these ways.

It seems to me that Kratzer has failed to come to grips with the thought that the modal auxiliaries are ambiguous, and that her arguments are rather unconvincing. Firstly, we should dissociate the ambiguity view from the implausible 'Aristotelean' move. On the face of it, this move involves a confusion of the grounds for an assertion with its conditions of truth. To take an analogy, no one would claim that the sentence

(23) You are overweight

meant, in a particular context, the same as

(24) In view of your excessive eating for the last month,
you are overweight,

even if 24) spells out the speaker's reasons for asserting 23). For 23) could be true if, say, the person had been injected with too much of a special fattening chemical, whereas 24) could not. Similarly, in a context where it is presupposed that a certain person is incapable of begetting through immaturity, "He cannot beget", construed as "In view of his immaturity, he cannot beget", would have to be taken to be false if the incapacity was caused in some other way.

I take it that this consequence is unacceptable enough for us to reject this move.

The ambiguity theory should only recognise a handful of senses of can and must - our categories A, C and D, perhaps. Sometimes we can tell the kind of modality being expressed by the tense of the verb. must have with the past participle is reserved for $\Box P$, had to with infinitive expresses $P\Box$. This should be borne in mind when considering Kratzer's second line of thought, that the "different occurrences of the word must ... [in 21) and 22)] seem to have the same meaning." This seems to me to be a highly debatable point. must have does not lose its epistemic sense, even when embedded the way it is in 22). If we rewrite 21) using must have

(25) In view of what their tribal duties used to be, the Maori children must have learnt the names of their ancestors,

we have an epistemic modality, approximately: "we conclude that they learnt the names". Changing must have learnt to had to learn in 25) produces past obligation, intuitively - not a neutral must. Kratzer has given no thought to the question of how her neutral must - let's use \Box for it - would interact with tense. What could the significance of $\Box P$ and $P\Box$ be (if not $\Box P$ and $P\Box$)? As is suggested by our previous discussion of 23) and 24), the role of an in view of phrase is to qualify the main clause, not to extract part of its meaning.

The ambiguity theorist need not make the absurd denial of there being no common element to the various kinds of must he discerns - clearly they are all different kinds of necessity. His point is simply that there are a few

significant differences between them. It may be felt, however, that the discussion of Aristotle's example does not show anything wrong with the relational view. Perhaps the wrong relativisations were chosen. If we expand it in a different way, "In view of the way he is constituted, he cannot beget", we do not get the counterintuitive conflict in truth values. Moreover, there are more plausible examples of a fluctuation in sense, in the way predicted on the relational view. One such is provided by David Lewis (in [64]), who cites Kratzer's account with approval. Lewis envisages himself arguing with an elected official over ways to handle a potential embarrassment for the latter. They have been discussing options, such as destroying the evidence, which are available within the constraint - implicitly assumed - that the official does not commit political suicide. Lewis rudely interjects: "There is one other possibility - you can put the public interest first for once!" That would be false, ie. the possibility in question would not be a genuine one, if that constraint, determining what is and is not to count as possible, remains fixed. But Lewis's point is that his very utterance protects itself against this by widening that tacit delimitation of the possible, in accordance with a rule of accommodation. A rule of accommodation is in effect a principle of charity: a rule for changing the 'conversational score' in a context, eg. the presuppositions of the participants. Some of the cases Lewis cites shift the score back to a previous point as easily as it was shifted away. For example, salience with respect to an elliptical the cat can reinstate one cat as the topic of conversation after some other cat had usurped that position. But in other cases he maintains that

accomodation tends to work in one direction only. In this example, this means that once the boundary has been widened, it remains there and is difficult to shift back to the previous position of excluding Lewis's possibility. If, after Lewis's interjection, the official protests

(26) I cannot do that

according to Lewis he is mistaken.

It is not clear to me that Lewis is right on this last point. We can understand why the official would make that protestation. The point about these practical can judgements is that certain consequences follow from their acceptance. If 26) is accepted, the person in question is automatically exempted from doing the action. If "You can do that" is accepted, it is clear that the person has a choice, and is responsible for whether or not he does the deed. If the official did assert 26), he would be attempting to shift the boundary back to where it lay before Lewis's interjection. The suggestion would be that staying in office is to constrain the consideration of options. It seems as implausible in his case as it is in Lewis's to say that his assertion would be false. The official is a normal speaker and so will have an implicit mastery of the 'kinematics of conversational score'. If the boundary stays fixed at the outer limit, after Lewis's interjection, he should know that. So his assertion of 26) would require him to have the belief that he cannot put the public interest first, even if he is prepared to commit political suicide. And this is very implausible. The dispute between the two is unlikely to be over this kind of fact - they can both agree that he cannot if he is not prepared to give up his office, and that he can if he is.

According to Lewis, "the boundary between the relevant possibilities and the ignored ones ... enters into the truth conditions of sentences with can ..." If we represent Lewis's belief that the official can put the public interest first as $(\exists w') \text{PRT}_1(@, w') \& \mathcal{S}(w') \varphi(a)$, then the official's protestation 26) will either represent the negation of this, or involve a more restrictive accessibility relation: $\neg(\exists w') \text{PRT}_2(@, w') \& \mathcal{S}(w') \varphi(a)$. But neither of these alternatives is plausible. The first requires the official to have a false belief, and the second makes it impossible to see how there is a dispute between them, since both claims are acceptable to both of them. The negation test strongly suggests that the drawing of a boundary is a matter of presupposition. Any ordinary judgement about whether something is possible presupposes that some boundary between possible and impossible has been fixed. But that boundary is not strictly a part of the sense of the statement⁶. What may shift, in a context, are the background assumptions about the delimitation of possibility - conditions of truth, strictly speaking, remain the same.

We must therefore conclude that the ambiguity theorist has a more accurate description of the facts than the relational theorist. Lexical ambiguity, Montague suggested, would be accounted for by rejection of the assumption that there is just one actual model, a unique actual dictionary. However, in view of the fact that there are ambiguous words of special interest whose interpretations could be fixed, it would be preferable to adopt Bennett's proposal and have the required number of distinct lexical items, each tagged by an index which the syntax would be designed to erase.

4: THE PROBLEM

In Chapter 1 I observed that it would be a defect of a truth-theoretic treatment of modality if it could not be located within, or at least in relation to, a general semantic account of intensional contexts. In this section I want to consider the difficulty as it faces model theory based on possible worlds. In a nutshell, the problem is that equivalence of extension in every possible world does not seem sufficient for synonymy, and so as Dummett has put it "cannot be meant to give a representation of the knowledge that someone has when he knows some expression in a language" ([33], p.422).

The two versions of the problem which are most discussed concern sentences and proper names. I shall introduce it via the first version, since Montague's extant thoughts on it relate to that; I shall then go on to discuss the second in more detail, with special reference to the writings of Saul Kripke. It will be convenient if we begin by restating, in simplified form, the relevant interpretative axiom for believe. It is based on the idea that Maggie believes that φ iff φ is true in every possible world which is consistent with her beliefs:

$$(1) \quad T(\alpha \text{ believes that } \varphi, w) = (\forall w') (\text{Belief}(R(\alpha), w') \rightarrow T(\varphi, w')).$$

The ML thus employs the predicate " $\text{Belief}(\text{Maggie}, w)$ ", which has as its extension the set of Maggie's 'belief worlds' - the proposition which encapsulates the totality of her beliefs. Maggie believes that φ iff that proposition is included in the proposition expressed by φ . Then any sentence ψ logically equivalent to φ , ie. one for which the

following will be true:

$$(2) (\forall w)(T(\varphi, w) \equiv T(\psi, w))$$

must also be believed by Maggie.

Montague contemplated two responses to this result. One was to accept it as revealing a genuine feature of any properly proposition-oriented notion of belief. Its acceptability, he claimed, is bolstered by taking seriously assent as a criterion of belief ([73], p.139) - basically¹ what Kripke has termed the "disquotational principle": "If a normal English speaker, on reflection, sincerely assents to 'p', then he believes that p" ([62], pp.248-49). There is something to this. If Maggie sincerely assents to both φ and to $\neg\psi$, where ψ is a logical consequence of φ , there is clearly a real problem for us about what beliefs to ascribe to her. Does she really believe the proposition expressed by φ ? We have as much right to ascribe to her that belief as the belief that $\neg\varphi$, given that we have as much right to ascribe to her the belief that $\neg\psi$, and given that $\neg\psi$ logically entails $\neg\varphi$. One may, then, maintain that if someone really believes some proposition, they do believe all its consequences, implicitly at least. Why, after all, should we be aware of all our beliefs? But even if we may distinguish this sense of belief, there will still be other propositional attitudes - be aware that is an example Thomason cites - and therefore, we may assume, a different sense of belief, for which the above result is unacceptable. We should not forget that the case just considered was a rather special one - the general problem of logical fallibility is not that one explicitly assents to the negation of some consequence of a belief, but that being unaware of all the consequences one simply doesn't know whether to affirm

or deny them. This suggested to Montague at one stage that we are dealing here with more linguistically-oriented notions ([73], p.139). Obviously if a context (eg. "... is an instance of L.E.M.") is implicitly quotational, we cannot expect it to permit intersubstitution of logical equivalents salva veritate. The treatment which Montague actually adopted for these notions which do not permit interchange of logical equivalents was to allow that even if

$$(2') T_*(\varphi, g, w, t, \mathcal{M}) \equiv T_*(\psi, g, w, t, \mathcal{M})$$

holds for all g, \mathcal{M} and all designated points of reference $\langle w, t \rangle$ ($= \varphi$ and ψ are logically equivalent), there could be certain "unactualisable" points $\langle w, t \rangle$ for which 2') fails ($= \varphi$ and ψ are not synonymous). This proposal has the air of being a purely technical solution. For one thing, when Montague made it he still affirmed allegiance to "the best and most elegant approach ... viz. to permit unrestricted interchange on the basis of logical equivalence ..." ([73], p.231). More importantly, there remains the question of the status of these seemingly 'impossible worlds', the invocation of which might well be thought to bring the whole possible worlds framework into disrepute. (At non-designated points logical constants are assigned deviant extensions). Given the unclarity of these things, it remains unclear what explanation has been offered either of synonymy or of the propositional attitudes we are interested in (especially if they are supposed to be oriented towards linguistic entities).

To see whether there is any plausibility in this move, it will help if we turn to consider the second version of the problem. Imagine someone coming across entries [53] and [54] in the bibliography. He may wonder ^{whether} Hans Kamp and

J.A.W. Kamp are two different people or not. The following thought might enter his mind:

(3) It may be that Hans Kamp is not J.A.W. Kamp.

Other epistemic modal operators suitable for use here would be maybe and Kripke's favourite it could turn out that. On both the description and causal accounts of reference there need be no cause for complaint that the person is less than competent in the use of the names. For the description theorist, he can associate properly identificatory descriptions with each name, eg. "the author of 'Two Theories About Adjectives'" with the latter. For the causal theorist he has picked the names up from a speaker standing in an appropriate causal relation to the bearer of the names. I choose the example because it is a realistic case, closer to home than ancient Babylon, of a question arising as to whether or not two names stand for the same entity. We know the kind of situation which the person would envisage, or at least, could not exclude as a possibility - his coming to accumulate sufficient further information that he could associate incompatible properties with these two Kamps. Perhaps he imagines meeting them side by side at a conference. The problem is this. Since Hans Kamp and J.A.W. Kamp are one and the same person, that imagined situation is not a logical possibility - in no possible world can he be other than himself. So a first glance might lead one to think that, by turning to the subsentential level, we have found a vindication of Montague's tolerance towards logically impossible worlds. (At the sentential level, that impression might have been gleaned from an example like 2.1.15 - for there the truth actually ascribed to the whole propositional attitude construction should be a function of the set

of worlds in which the content sentence is true, yet that sentence is not true in any logically possible world.)

Whether this is a more than superficial diagnosis depends in part on whether or not ordinary proper names are rigid designators. The two main theories - or pictures - of naming tend toward different descriptions of the situation, neither of which is very palatable. On the description theory, one might assume, names are not rigid designators, being synonymous with some associated description. Hence there need be no impossibility about the circumstance envisaged by the utterer of 3) - it will just be one with respect to which the two names refer to different objects. But the description theory would appear to face the following dilemma. Either the explanation is offered at the community-wide level, or it is a question of descriptions being associated with names by individual speakers. The former seems straightforwardly false, since for example most of those who are competent with these names do not attach any difference in sense between them. If it is the latter then, as Kripke has pointed out ([62], p.245), the sense which I attach to the names, apparently being the same for both, cannot be sufficient for me to render intelligible the epistemic state of the person concerned by anything I can say using the names, eg. "He believes that Hans Kamp may be other than J.A.W. Kamp". For I would be attributing to him a belief in the epistemic possibility of the denial of a trivial self-identity. But the causal theory would appear to fare no better. If names rigidly designate their bearers the truth of 3) should require an epistemically possible world with respect to which the two names rigidly designate a non-self-identical object. But this is nonsense. It does

not help to render intelligible a seemingly impossible belief to import a contradiction into our description of the case.

Let us turn to the writings of Kripke and Dummett for some enlightenment on these matters. They are agreed at least in that each has, at some point, called talk of possible worlds "metaphorical" ([60],p.80, [29],p.127). For Dummett, since the rigid/non-rigid distinction is framed in terms of possible worlds, when we give it non-metaphorical content it reduces to the distinction between a term's taking wide or narrow scope in relation to a modal operator. The rigidity of proper names becomes the thesis that their scope is always wider than that of the relevant modal operator. A related point is made by Kamp in connection with temporal discourse, in likening proper names to the now operator (names "refer back to the moment of assertion: they point at the object they name at the time the assertion is made..." [55],p.187) - since now always takes widest possible scope (cp. §1.5). But in his new Introduction to [59], Kripke explicitly rejects the relative scope interpretation, on the grounds that it cannot be as general as the rigidity doctrine, requiring as it does the presence of modal (or tense) operators. The question of rigidity arises with respect to ordinary understanding of non-modal sentences, since understanding these requires knowing what has to be the case for their truth - ie. whether or not such a sentence would be true in other possible situations. In the context of a dispute where it is common ground that possible worlds are metaphorical, this may seem an odd reply - Dummett's purpose was to remove talk of those things. How better to do this than to reintroduce modal operators?

The whole point of introducing possible worlds was originally to account for the truth conditions of sentences involving operators - ordinarily one would say we have no need to invoke them in giving the semantics for a non-intensional (part of a) language. Kripke might deny that - he might claim that even in an extensional language, names and other terms require mention of modal facts concerning them if a full semantic characterisation is to be given. But even if we grant him what would be for the purposes of a truth definition (of either of the two kinds we have considered) strictly superfluous, there remains the question of what relation these modal facts bear to the semantics of the modal part of the language. If we simply consider the ^{non-modal} ~~non-modal~~ reference of a term occurring in a non-modal sentence, in relation to some unactualised world, we will be in danger of losing touch with the very linguistic phenomena we originally sought to explain. This is perhaps one reason behind Kripke's continuing reluctance to admit that descriptions can be used rigidly. If we set things up by asking after the truth value of a sentence like "The last great philosopher of antiquity was fond of dogs" in relation to some imagined counterfactual situation, it will only be natural to consider who is imagined to be the last great philosopher of antiquity in that situation, and thence to assume that a description will always function non-rigidly. While it undoubtedly follows from my understanding of a sentence that, had things turned out differently, I could ascertain its truth value in relation to that circumstance, we must be clear on the precise significance of this fact. A hypothesis about rigid or non-rigid designation, whatever its intuitive appeal (or that of the correlative metaphysical thesis), is

only as good as the phenomena it seeks to explain. In the context of the present dispute, the methodology just described is simply question-begging, since granted the equivalence $T(\Diamond\varphi, w) \equiv (\exists w')T(\varphi, w')$, it amounts to the assumption that the possibility operator is always to be semantically evaluated prior to the description, ie. that descriptions always fall within the scope of the modality.

If a rigid designator is, in possible worlds terms, one which with respect to any possible world refers to the same entity as it does in actuality, then the acid test for a definite description's being rigid (or more precisely, being used as a rigid designator) should be its use to pick out the individual which it does denote even with respect to a world in which some other entity is the unique satisfier of the description. In scope terms, then, we should listen with our native's ear to a sentence like this

- (4) The 37-th president of the U.S. need not have been the 37-th president of the U.S.

to hear if one or other occurrence of the description has wider scope than the modality. On the natural reading of 4), the first occurrence does - the sentence might be paraphrased as conveying that Nixon was only contingently the 37-th president of the U.S. This has little to do with the description occurring prior to the modal expression in the sentence. The scope we assign to a description waits upon a comprehension of the whole sentence in which it occurs. If we convert the second description into an explicitly rigid term (of a kind which Kripke acknowledges, [59], fn. 22):

- (5) The 37-th president of the U.S. need not have been the actual 37-th president of the U.S.

the first description intuitively has narrow scope.

Roughly, 5) means that we could have had a 37-th president of the U.S. different to Nixon. With

(6) It need not have been that the 37-th president was the 37-th president,

we have a sentence ambiguous between a reading like 4)'s and the false one which gives both descriptions narrow scope - false in that, however things had turned out, the 37-th president would have been the 37-th president.

On the current understanding of the causal theory, therefore, proper names take scope wider than modal expressions (what is distinctive about them is that, unlike descriptions, they always do). We might thus rewrite 3) as:

(7) $(\lambda v_1)(\lambda v_2)\text{Maybe}(v_1 \neq v_2)[\text{Hans Kamp}, \text{J.A.W.Kamp}]$,
showing clearly that there should be no difference between 3) and a similar sentence using either of the two names twice over. The same difficulty is brought out on Kaplan's 'direct reference' formulation of the theory, according to which when a proper name occurs in a sentence, the object itself which is the bearer of the name is the relevant constituent of the proposition expressed. That would therefore make the content of 3) quite mysterious, the same object being involved in it twice over.

One approach to the problem posed by 3) and its ilk would be to try to qualify the theory in some way. The idea would be that although 'direct' or 'simple' reference must correctly characterise names in essence, there could be certain peripheral cases where names are used non-rigidly. The difficulty facing any such modification is evident from the old description theory: what other object, and in virtue of its possessing what property, could a name refer to with respect to another world if not its actual referent? There

is a much-maligned though to my mind plausible candidate - with a name such as Lauben it would be: the thing called Lauben. To appreciate its plausibility we might start with the way we typically become familiar with previously unencountered proper names. A normal speaker of English who has, inter alia, mastered the business of referring to things by name, has the ability to add endless new names to his lexicon. He can generally recognise an expression as a name on the basis of typography; he can then form the belief that it has a referent - an inference from previous cases. If someone tells him "Dr. Lauben has been wounded", when he has never heard the name Lauben before, then the natural view, which I think is correct, is that he will at least have understood that someone called Lauben was asserted to have been wounded. He can use a name correctly, having only heard or read it once, by intending to refer to the thing which, according to the usage on the occasion of his encountering it, it is normally taken to refer to. Perhaps he comes across the name on an otherwise blank piece of paper, but if he has identified it as a name, he can still say "Who is G. Lauben?" (or perhaps "Who or what ...?"). "What does 'Lauben' mean?" is an inappropriate question for a name, the asking of which suggests that the inquirer has not identified it as a name and which might therefore prompt a reply prefaced by the remark "Well, its a name ...". It is thus implicit in this bit of ordinary usage that a name has no meaning - we assume that it simply stands for a particular object, on encountering it, and what we want to know is which object (or at least some partially identificatory information about it).

No one could pretend that this can be the whole

story about reference, since it blatantly depends on the assumption that the name will already have its reference established. This does not preclude it from being an important chapter in that story, however. The fact that language use is generally accomplished unreflectively, even when mastering new names, should not be allowed to obscure the fact that, implicitly or explicitly, we form beliefs which are semantical in nature. Language is part of the world we experience. There is thus one definite description which a competent speaker at least implicitly believes to denote (relative to the context) the bearer of a name - "the object called 'Lauben'", in the case of Lauben. (Such a description, like "the postman", will be contextually variable because elliptical - but is still likely to be definite when used in a context.) Normally speakers only use names on the presumption that they have bearers (obviously this is subject to qualification concerning the use of names from fiction as if they had referents.) So Kripke is wrong to say that, on the description theory, people who know of Feynman and Gell-Mann only that they are leading contemporary theoretical physicists must attach the same sense to the names. The property of being someone called Feynman is not the same property as that of being someone called Gell-Mann. But Kripke would be right to say that if that's all there is to the description theory, it can't be enough. He brings out the incompleteness of such an account in his discussion of Glunk. One cannot determine the reference of that name by saying "I shall use the term 'Glunk' to refer to the man I call 'Glunk'". The circularity gets one nowhere. "One had better have some independent determination of the referent of 'Glunk'" ([59], p.286). This does

not tell against our assuming a comparable attitude when we acquire (as distinct from: try to introduce) a new name: "I shall use 'Glunk' to refer to the man who is called 'Glunk'". One passes the buck, not to the man from whom one first picked up the name (cp. [59], pp. 297-99), but to the community-wide determination of its reference, which presumably will trace back to some initial act of baptism.

At this point Kripke might muster an argument which is at least implicit² in his line of reasoning, against any association of a name with a description. It comes as part of another dilemma posed to the description theory: either it is providing an account of the sense of a name, in which case it is false, or it is providing an account of how its reference gets fixed, in which case it will be inadequate³. The part we are concerned with is the first, which proceeds by drawing out unacceptable metaphysical consequences (certain properties being essential to an individual, eg. Nixon) from correlative semantical claims (the sense of Nixon). It might be thought that α need not have been β (or might not) is a general recipe which, when truly instantiated, shows that i) the referent of α possesses the property expressed by being β only contingently and that ii) the sense of β cannot be part of the sense of the term α . Whereas i) almost uncontroversially follows from the meaning of "contingent", ii) is far more dubious. Its weakness should be apparent from our discussion of 4) above. The fact that the 37-th president was only contingently the 37-th president could hardly show that "the 37-th president" differs in sense from "the 37-th president"! All normal definite descriptions express a property which is only contingently possessed by the object which, if any, they

denote. So the fact that a description only expresses a contingent property of an object does not by itself show that it must differ in sense from some codesignative name.

Perhaps the point Kripke was most concerned to make in this connection was as follows: that if it comes to a showdown between a name and a description, over which should refer to the thing which they both actually designate, with respect to some other possible world, it is the name which will always win out. That is to say, if we consider a sentence like

(8) G. Lauben need not have been the person called

'G. Lauben',

it is the description which is forced to let go of Lauben himself, rather than the name. Now while 8) undoubtedly has a contradictory reading which gives the description wider scope, we can also take it this way - what is more naturally expressed by (but with which 8) is not to be confused) a sentence without a definite description:

(9) G. Lauben need not have been called 'G. Lauben'.

But for this to be significant in the present context it would have to be the case that, faced with an assertion of (10) G. Lauben need not have been G. Lauben we would not construe that too along the lines of 9). As it happens, that is the way native speakers tend to read it.

Kripke is apparently prepared for this kind of thought: "Sloppy, colloquial speech, which often confuses use and mention, may, of course, express the fact that someone might ... not have been called 'Aristotle' by saying that he might ... not have been Aristotle. ... Colloquialisms like these seem to me to create as little problem for

my theses as the success of the 'Impossible Missions Force' creates for the modal law that the impossible does not happen" ([59],fn.25). But is it so simple? The acceptance of 10) (or 3)) as true by ordinary speakers is not a challenge to the modal principle that any objects x and y which are identical are necessarily identical. We can agree that the planet Hesperus cannot be other than Phosphorus, which is to say that in one sense of "necessary" - the philosopher's logical/metaphysical necessity - "it is necessary that Hesperus is Phosphorus" is true. The question is (as Kripke points out in his new Introduction) whether an ordinary proper name like Phosphorus (if it is one) refers to the thing we actually call Phosphorus with respect to every world. Consider Kripke's discussion of (11) Hesperus might have turned out not to be Phosphorus. According to him this way of speaking is "loose and inaccurate" ([59],p.353). It should be clear from the account of 3) that he is mistaken about this - 11) is a past tense version of a sentence which, like 3), may be a natural, almost unavoidable form of words for those concerned. Moreover, the circumstance which Kripke describes ([59], pp. 306ff.) in which another planet such as Mars occupies the morning position of Venus and gets named Phosphorus, although it wouldn't be a situation in which Hesperus wasn't Phosphorus, ie. wasn't itself, would nevertheless be the kind of circumstance one would ordinarily cite in defense of 11).

It seems to me that the Kripkean arguments we have considered do not provide us with any convincing reason why the sense of a name α cannot be equated with that of a description the thing called α though, to repeat, he has

shown that there is more to be said than this. It is significant that in his discussion of names in belief contexts Kripke stresses both that this is an area where our normal apparatus for attributing belief, based mainly on the disquotational principle, appears to break down, and also that there is no problem for our description of the puzzling cases as long as we remain within the essentially linguistic terminology of the thing called α , rather than using the name α itself. In the case of the itinerant Pierre ([62], section II), we can say that he believes that the thing called Londres is pretty, and that the thing called London is not - but can we answer the question: does he, or does he not, believe that London is pretty? ([62], p.259). There seem to be two ways of approaching this. One is to decline the terms of the question, ie. to refuse to apply the disquotational principle to Pierre's sincere assent to "Londres est jolie" and "London is not pretty". The specification of the content of Pierre's beliefs remains in the explicitly metalinguistic form already recorded. We thus retain the link between London and Londres at the expense of that between London and the thing called 'London' etc. The other response is to reverse the connections, treating Londres and the thing called 'Londres' as intersubstitutable, and replying that Pierre believes that Londres is pretty but that London is not. This is a natural way of describing the situation and one which keeps the linguistic element of Pierre's beliefs implicit. Since the thing called 'London' and the thing called 'Londres' are not intersubstitutable, we are entitled to say, on this response, that Pierre does not believe Londres is London without fear of contradiction, just as we would want to say

a similar thing in the two Kamps case. This approach therefore coheres best with our need to say things like 3).

We should conclude, it seems, that Montague was wrong to suggest that we have a choice in whether to construe belief as proposition- rather than language-oriented, and thus whether or not we accept that logical equivalents are intersubstitutable in all intensional contexts. The traditional notion of a proposition is that of objectively communicable and graspable content, lying behind the linguistic apparel of the sentences used to express it (the sartorial imagery goes back at least to "Der Gedanke"). In the totality of genuinely possible worlds, of objective conditions of truth, we reach the limits of this notion, realistically construed. What is needed in order to handle belief contexts in general is not a dabbling in impossible worlds, or a subjectivisation of the notion of sense - for that way lies incoherence or incommunicability - but a recognition that, for certain cases at least, the linguistic garb is an ineliminable part of what is expressed, ie. cannot be removed, as the disquotational principle enjoins us, in the move from sentence to proposition expressed. It may be that model theory, with such techniques as the discourse representation structures of Kamp, has now the means to produce an adequate treatment of the problem.

1: INTRODUCTION

There are many facets of Michael Dummett's writings on semantics. I shall concentrate on two central themes, which relate to his critique of unthinking acceptance of realism: the argument for some kind of verificationist theory of meaning, and the thesis that metaphysics is subservient in some sense to the more fundamental discipline of semantics. The arguments will be developed as much as possible with reference to modal locutions. As Russell pointed out ([93],p.259), the verificationist has a particular theoretical interest in elucidating modality, since we need to know eg. what notion of possibility is involved when something is said to verifiable, ie. such that it can be verified. According to Dummett, consideration of vagueness shows that the idea that verification be possible in practice leads to incoherence - from which it follows that if the demand for verifiability is a worthy one, it must be shown not to lead to the strong demand of practical verifiability.

For Dummett, the emphasis on a theory of meaning is that it is a theory of understanding. To understand an expression, ie. to grasp its meaning, is to know how to use it, to know what its role in the language is. As yet this is to say very little; it serves to remind us that any account of meaning must square with the principle "Meaning is use". Various interpretations of this Wittgensteinian dictum are possible (cp.[29],pp.359ff.). Taken to extremes

- "Don't ask for the meaning,..." - it embodies the rejection of the idea of a central core of sense which can be extracted from the motley of our linguistic interactions, and therefore of a systematic theory of meaning which aims to isolate it. Less radically, it tolerates the notion of sense as merely a convenient shorthand for summing up facts about usage. This interpretation seems the one which harmonises best with Wittgenstein's own account of understanding. We are not to conceive of sense as if grasping the sense of an expression is something that can be achieved independently of, and therefore having the potential to conflict with, a mastery of its use. For example, as Kripke and others have persuasively argued, to be able to use a proper name correctly does not require one to be in possession of a nontrivial criterion for identifying its bearer - so requiring that the sense of a name supplies such a criterion would violate this version of the principle. Construed thus, meaning is not something which could be made out by a theory in a way that would show up as mistaken certain linguistic practices, such as the use of L.E.M.

As Dummett understands it, the dictum breaks into two halves: "meaning determines use", and "use determines meaning" (cp.[33],p.216). On the one hand, the meaning of a sentence doesn't simply consist in the different uses to which it can be put, but on the other, we cannot afford to sever the connection between the two. This ambidextrous interpretation embodies the Fregean distinction of sense and force - in terms of a theory of meaning, its division into two subtheories, a theory of sense, associating a specific sense with each sentence in virtue of its semantic

constitution, and a theory of force, containing general principles which link the sense of sentences with their uses (cp.[36],pp.12-13).

The problem a realist truth-conditional meaning theory faces concerns its ability to connect up sense and usage in this way. The meaning which resides in our language is brought into being by humans engaging in linguistic activity, it cannot therefore be made out to be something humanly incommunicable. It must be capable of being manifested in the use we put sentences to, and of being acquired from such use. A meaning cooked up by a theory which a person was incapable of making manifest would also be something no one else would be capable of acquiring from them. Thus conceived, the principle is aimed primarily at those who would make meaning ineffably private, making it something more than might ever come out in use (cp.[41],pp. 35-36,[85],p.8). But in the hands of the anti-realist, it is aimed specifically at theories which indulge in epistemically transcendent conceptions of meaning. The theorist must look to the conditions under which speakers are taught to accept and assert sentences as true or false. "What we learn to do," writes Dummett, "is to accept ... the occurrence of certain conditions which we have been trained to recognise, as conclusively justifying the assertion of a given statement ..., and ... the occurrence of certain other conditions, as conclusively justifying its denial" ([33],p.362).

The import of this consideration is itself open to interpretation. It might be taken to suggest that we learn a language by associating particular circumstances with particular sentences - those we are explicitly taught as

being conditions of its truth, or falsity. The problem would then be how we could understand previously unencountered sentences, and even those we have met, as true or false in circumstances we did not encounter in training, and therefore weren't taught whether to assimilate to the conditions of truth, or of falsity. We could not assume a notion of truth going beyond what has already been determined as such. Although an anti-realist may want to say that of course we go beyond the limited experience of our linguistic training (but not as far as the realist would have it), we shall see that he is sometimes dogged by this strong interpretation of "use determines meaning". But he would rather put his case as follows. For both decidable and undecidable sentences there will be possible situations in which we would recognise them as true, and others in which we would recognise them as false - if they occurred and we encountered them. With a sentence of either kind, our current epistemic situation may be neither verifying nor falsifying. The point about decidability is that with such a sentence we are entitled to conceive ourselves as understanding it bivalently, because it is ensured that we can gain access to the situation in which its truth value is recognisable. The realist, of course, wants to say that if our conceptions thus extend beyond what we actually experience, in the decidable case, then why not also for undecidable statements? To him it looks like the same confusion of meaning with evidence in both cases - the anti-realist who concedes decidable sentences to him is making the same mistake, with regard to undecidable sentences, as the anti-realist who doesn't.

At this point the dialectic can be illuminated by

consideration of the anti-realist's principle that truth entails knowability/verifiability/provability:

$$(1) \quad \varphi \rightarrow \Diamond K\varphi.$$

One reason why this principle may provoke hostility is a misunderstanding of it as reviving the old verificationist spirit of 'the method is the message'. If we had to possess a method of verification - one guaranteed to result in the determination of the truth value - in order for a statement to be meaningful, undecidable ones would be meaningless. And surely, for example, the jungle linguist could traverse the Amazon and lose his front door key in its midst, so that any search we might make for it would have no guarantee of success? Quite so, the anti-realist would agree, but what would it be for the key to be there and also that it be impossible for it to turn up, even by chance? That would be a model transcendent truth condition, transcending all possible evidence - something not in question in this case, since it must be possible for us to hap upon it, just because its there.

Even so, it is likely that the realist will regard 1) as either vacuous or false. In the sense in which he finds it acceptable, 'knowability' is such a weak notion that it need bear little resemblance to our own powers of acquiring knowledge. The anti-realist has just granted that a situation is knowable just because it exists, and what exists, the realist stresses, does not depend on our epistemic predicament. Alternatively, if the notion of verifiability is made more concrete and related to our powers, 1) is false. This dilemma he sees the anti-realist facing is sharpened by a proof, recently resurrected by Hart and McGinn ([49]), which shows that 1) entails the stronger

principle

(2) $\varphi \rightarrow K\varphi$.

2) is proved by reductio, ie. by first assuming its negation - equivalently $\varphi \ \& \ \neg K\varphi$ - and deriving by means of 1) and some uncontentious principles, the conclusion $\Diamond(K\varphi \ \& \ \neg K\varphi)$ which, unacceptably, asserts the possibility of a contradiction. This highlights the vacuity of the vacuous interpretation. If all facts are knowable simply because they exist, as part of reality, then already the idea of a knowing subject plays no essential role in the notion of knowability¹. At most one need assume a hypothetical omniscient being, one who would know the facts, if he existed. So we might as well assume that all facts are 'known' by it, ie. would be known, if he existed, simply because they make up reality. It also rubs in the falsity of 1) when knowledge is given a more mundane interpretation. Clearly, we can't know both that something is the case and that we don't know it. So, since we can't know it, by 1) it can't be true. So either the upholder of 1) is forced into the more radical anti-realism which disallows a notion of truth beyond what is known, or he must temper the generality of 1). A moderate anti-realist would say that 1) needs restricting, to cope with knowledge of knowledge.

At the heart of the realist's antipathy to 1) is the thought that a state of affairs can exist, yet there be all kinds of obstacles to prevent us discovering it so - even by chance. He thus allows for the metaphysical possibility of there being states which exist in total isolation, epistemically, from us. For the anti-realist, however, when evaluating 1) we must bear in mind that truth is a relation between two terms: the world, and language -

our language. The realist begs the question in the crucial cases by assuming that the world already is determinate. It is as if he could see beyond the obstacles, mentally moving them aside, then replacing them at will. Take the oft used example of infinite quantification. We gain our understanding of the quantifiers with finite samples. In certain cases, we observe something of the form $(\forall v)(\varphi(v) \rightarrow \psi(v))$ to be true in the most direct manner possible, collectively. We can observe all the φ 's and see whether they are all ψ , as perhaps with "All your fingernails are red", with the hands held close together. More frequently we have to adopt a less direct approach, checking through each φ to see whether or not it is ψ , and finally ensuring that we have checked them all (unless already encountering a falsifying instance). By thus running through the instances we can ascertain whether "All the rooms in this house have fireplaces" is true. In both these cases we can imagine getting indirect evidence - eg. for the last one, from the architect's plan. But there are also cases where indirect evidence may not exist or at best be inconclusive. Thurber once introduced the category of the "fascinating but undemonstrable" generalisation, into which he put "People who break into houses don't drink wine". The 'undemonstrability' in this case relates to the enormous effort that would be required in practice to check through every contemporary american housebreaker (assuming this to be the intended domain), coupled with the inconclusiveness of any indirect evidence. It is not, if present tensed, in principle undecidability. The obstacles to verifying it are ones of degree. Since the original domain was assumed finite, by restricting it - eg. "The

housebreakers in this town ..." - we could make the task more manageable without changing its essential nature. Thurber's generalisation is 'in principle' verifiable, like the key in the jungle example - the difficulty is quantitative rather than qualitative. Not so in the infinite case, for here the obstacles are not ones we can surmount, or remove without changing things radically. If someone understands a sentence involving such quantification, he should be able to acknowledge a demonstration of it if presented with one - but we cannot take seriously the idea that we are guaranteed such a demonstration, or failing that, a refutation. We would only have this guarantee if we could check through an infinity of objects, whereas all we can do is check a finite subset of any such set - not the same thing at all. It is here that the realist abuses the license gained in the finite case and pictures the condition for the truth of a quantified sentence as obtaining independently of what we, in conferring meaning upon it, can take as establishing its truth. A non-(effectively-) decidable sentence is one for which we have no effective method for determining the truth value. In the undecidable case, to say that a sentence is bivalent is to invoke a notion of truth which cannot be reconciled with our ability to use it - to recognise it as true if true, or false if false.

We can summarise the situation so far by considering how Dummett and Quine would juggle with the following theses about meaning and truth². a) The ultimate ground of the meanings of sentences - the source of our acquaintance with them - are the circumstances in which their assertion would be justified (where "circumstances" must

be understood liberally - cp.[34]p.132). b) If and only if truth conditions are completely determined, meanings are³. c) Truth is not an epistemic notion - it may transcend our ability to recognise it. d) Meaning is complete, ie. determinate. Quine and Dummett can be said to agree more or less on a) and b) (hence the comparisons). But then we must acknowledge the existence of sentences which are not effectively decidable, ones for which, by definition, we have no guarantee that we shall be able to recognise as true or false. Quine, realist that he is, accepts c) and so is forced to give up the determinacy of sense. The anti-realist, by contrast, cannot accept such anti-realism about meaning itself: meaning is something created by us and so must be totally accessible to us. Matters pertaining to sense, such as synonymy, must be effectively decidable (cp.[29],p.632). So he resolves the tension by admitting d), and advocates 'epistemologising' the notion of truth. By rejecting c), he thus transforms the nature of b), since meaning will not be given in terms of classical truth conditions (see here eg.[36],p.13). We drop the idea that truth conditions could obtain totally unrecognisably. The traditional realist slinks off, privately cherishing b), c), and d) at the expense of a), in violation of the usage constraint.

2: DISPOSITIONS

Having sketched the argument for rejecting the realist conception of meaning as bivalent truth conditions, I now turn to consider it unleashed against a particular instance of realism - that with respect to the meaning of disposition terms. These relate suitably to my general project, since to ascribe a disposition is "to make a statement about potentiality" (Goodman). Verifiability is one already mentioned, fragility is another. We say that something fragile can break easily, or can be broken easily. In this disjunction we see the intimation of a distinction between active and passive dispositions. But it is little more than grammatical (cp.[65],pp.99-100); most dispositions can be regarded as passive if only because something has to happen to their possessors to produce their manifestation. And for this reason many more properties can be regarded as dispositional than just the familiar ible's and able's. Any property expressing a state of a thing can be regarded as at least partially dispositional, for any state can be regarded as disposing whatever is in that state to react in certain ways. As Popper points out ([83],p.151), not only solubility, but also being dissolved is a dispositional property of a solute. As opposed to having reacted with a liquid, being dissolved requires a substance to be recoverable in a certain way - a disposition. But dissolving is not; it is not the disposition to behave in a certain way, but the behaviour itself. Though even here, as Goodman indicates ([44],fn.7,p.41), the distinction is not absolute. A phenomenalist might regard dissolves as a dispositional

term for certain kinds of appearances. The important thing is that we spell out the dispositional nature of a property in terms which are, relatively speaking, non-dispositional - paradigmatically, these describe features of observable events. Hence the elucidation of dispositional predicates by means of conditionals linking the manifestations - what it disposes the thing to do - and the circumstances required to produce these effects. The dispositionality of the redness of a surface consists in the fact that it either does or would look red, if you do/did the appropriate thing to it: if you observe it in normal light. The conditional may succeed in explaining the predicate in different terms, but as this last example shows, it may not. In the latter case, what elucidation is achieved comes through the description of the conditions in which it will be displayed. It is the realisation of these conditions which turn a can into a would (or at least a should).

Before continuing, I need to review some features of subjunctive conditionals. It will be helpful to discuss some logical principles which potentially govern them alongside their analogues for the metric future tense operator F_n . Thus we have L.E.M. applied to conditionals, and its future tense counterpart:

$$(1) (\varphi \Box \rightarrow \psi) \vee \neg(\varphi \Box \rightarrow \psi) \quad (1': F_n\psi \vee \neg F_n\psi)$$

What is called the law of conditional excluded middle is 2):

$$(2) (\varphi \Box \rightarrow \psi) \vee (\varphi \Box \rightarrow \neg\psi) \quad (2': F_n\psi \vee F_n\neg\psi)$$

In effect, it asserts that if a conditional is false, the opposite (same antecedent, contradictory consequent) is true. It excludes those 'middle' cases where we want to say that ψ might, or might not occur, given φ . It thus supplies one half of the equivalence:

$$(3) \neg(\varphi \Box \rightarrow \psi) \equiv (\varphi \Box \rightarrow \neg\psi) \quad (3': \neg F_n\psi \equiv F_n\neg\psi)$$

the other half, from right to left, being unproblematic except in the vacuous case (cp.[63],pp.79-80).

There are two different kinds of reason that might be given for the equation 3) of the internal and external negation of a conditional, which correspond to the transposition into the subjunctive key of two traditional views about the future, the Ockhamist and the Peircian. According to the Ockhamist, $F_n\psi$ is now true if ψ holds, n periods hence, in what in fact turns out to be the actual future - it doesn't yet have to be determined that ψ will hold then. On this understanding of the future tense, all of 1') to 3') come out valid. The Ockhamist finds his conception reflected in our ordinary talk of the truth of future tensed statements (cp. eg.[9],pp.160-61), so it would be unsurprising if we sometimes understood "it would be ..., if ..." in a manner comparable to "it will be ...": $\varphi \Box \rightarrow \psi$ as true if ψ would in fact have occurred, had φ . But there is a crucial difference between the indicative and subjunctive cases, in that there can be no correlate for the latter of our waiting to see whether ψ comes to pass, and thus of ratifying the truth of the prediction $F_n\psi$ even when made before the issue had been settled.

Consider the following thought of Ramsey's, recently cited by Peacocke ([81],p.60): "If we regarded the unfulfilled conditional as a fact we should have to suppose that any such statement as 'If he had shuffled the cards, he would have dealt himself the ace' has a clear sense, true or false, which is absurd." Peacocke accuses Ramsey of mistakenly rejecting an instance of 1) on the grounds that 2) is invalid in this case, by confusedly holding on to

the equivalence 3). But the Ockhamist reading is more charitable to Ramsey (and accords with what he goes on to say), For the Ockhamist, 3) does hold - as with 3') for Fn - but this time 2) fails, and so therefore does 1). Someone guided by this conception might protest "How can you say what would or would not have happened, had he shuffled the cards?" Since shuffling 'randomises' the outcome (assume, if you need, that a little free will would enter the performance), neither outcome is determined. But neither is there any waiting to see whether Ψ (dealer getting an ace) comes to pass since, in Ramsey's phrase, the conditional remains "unfulfilled". So there is nothing in virtue of which we would regard the conditional as either true or false.

Just because the Ockhamist conception has this effect of invalidating 1), there is a point in a stronger interpretation of the conditional which, as Peacocke intends, makes Ramsey's conditional assume a definite truth value, falsity. This exports to the subjunctive the Peircian view of the future, for which if $\text{Fn}\Psi$ is to be true now, it must already be determined that Ψ will occur at the appropriate time. Whereas for the Ockhamist, we need only consider the one possible future development of the world which turns out to be the actual one, for $\text{Fn}\Psi$, for the Peircian we need to consider all possible future developments, and be sure that Ψ holds in them all. Likewise for $\varphi \Box \rightarrow \psi$, it must be that ψ holds in all the possible φ -worlds which are most similar to the actual one. (It should be pointed out that these are not the only options available. On Stalnaker's theory, there is a unique most similar φ -world, giving the effect - illicitly on Ramsey's

view - of the Ockhamist idea of a single world at which to evaluate ψ , and thus validating 2).) On the transposed Peircian conception, the equivalence 3) cannot hold in general (cp. for the future, $\neg \text{Fn}\psi \rightarrow \text{Fn}\neg\psi$ fails), so the rejection of conditional excluded middle 2) does not necessitate reneging on the classical 1).

This brings us around to the second kind of reason which can be given in support of 3). It concerns the adoption of a realistic attitude towards some statement χ which grounds the truth of the conditional. We accept the equivalences

$$(4) \quad \chi \equiv \varphi \Box \rightarrow \psi$$

$$(5) \quad \neg\chi \equiv \varphi \Box \rightarrow \neg\psi$$

Note that 4) and 5) alone are enough to guarantee the equation of the internal and external negation of a conditional. (Take negations in 4).) Being realists with regard to χ , we also accept

$$(6) \quad \chi \vee \neg\chi$$

which, given 4) and 5), is sufficient for the truth of 1) and 2). This is the kind of case most favourable to the 'connection' account of conditionals, with χ indicating the connection between (the states of affairs described by) φ and ψ . In the special case we are interested in, χ is the ascription of a disposition, and φ and ψ describe how it is manifested. According to Dummett, it is not the relation between disposition and manifestation, as given by 4) and 5), which is in dispute for the realist and anti-realist. (It seems that the sense of would here must be Peircian. For consider an attempt to introduce a predicate when the outcome is not determined on fulfillment of the antecedent condition φ . To exploit Ramsey's

example: call a pack of cards "aceable" if one who shuffles it will deal himself an ace with it - ie. such that for an unshuffled pack, the truth of "It is aceable" reduces to that of Ramsey's conditional. The Ockhamist would have us believe that a never shuffled pack is neither aceable nor not aceable, and that a pack in use is aceable one moment, not so another - mainly the latter. This seems highly inappropriate - when introducing a predicate, we would like a sense of will or would which is false in such a case. Our natural reaction is to reject such a predicate on the grounds that no normal pack can be considered to be aceable.)

An example of 5), using a timeless is and an explicit temporal coordinate (dispositions can be acquired and lost), is

(7) Jones is not (water) soluble at $t \equiv$ If Jones had been immersed in water at t , he would not have dissolved.

Realism about solubility entails that if it's not the case that had Jones been immersed at t , he would have dissolved, then the right hand side of 7) must hold good. We could say: either the connection - solubility - is present, or it isn't, so one or other of the opposing conditionals must be true. Another way to put this is to say that we take a realist attitude to the test for solubility. Realism about testing amounts to holding that when we test for something, we test for something. Or, in Dummett's more refined formulation, we believe i) that if the test is applied (φ , in the above schematism), the state it reveals (ie. X , or $\neg X$, for this is what it is a test for), would have occurred anyhow, ie. independently of the performance of the test, and ii) that even if the test is not applied, there

is still some state of affairs obtaining such that if the test were carried out, it would reveal that actual state ([29], p.607). So whether or not φ , either χ , or $\neg\chi$. Both realist and anti-realist will agree to

$$(8) \quad \varphi \Box \rightarrow \psi \vee \neg\psi$$

but only the realist will go on to infer 2) from this¹. A realist attitude to the testing of some property, semantically ascended one stage, becomes a realist attitude to the testing of the truth, ie. verification, of the corresponding statement. It is a belief in objectivity: the independence of the truth of a statement, what is the case, from whether we desire or believe it, and more importantly in the present context, from whether we actually check for it. In mathematics, for example, it is intuitively plausible to regard proofs as the kind of thing in virtue of which its statements are true or false. Nevertheless we believe that mathematics is objective in that its statements, at least the decidable ones, are determinately true or false prior to our undertaking, if we do, to prove/refute them.

A realist attitude to the test for some property involves the belief that one of the relevant instances of $\varphi \Box \rightarrow \psi$ and $\varphi \Box \rightarrow \neg\psi$ is determinately true. As long as one is justified in maintaining bivalence for the corresponding χ , this will be acceptable, for one or other of the conditionals will be true in virtue of the presence or absence of the property, as recorded by χ . But few of us are likely to hold a realist conception of testing across the board, since few of us are wholesale realists. Renunciation of realism applied to a certain kind of statement, as envisaged here, will amount to a refusal to accept the

relevant instances of 6), 2) and 1). The anti-realist need have no quarrel with $\varphi \vee \neg\varphi$, $\psi \vee \neg\psi$, and hence 8), for it is $\Box \rightarrow$ which introduces the crucial element of undecidability. To see this, consider a not wildly implausible case of anti-realism, one by which Dummett has advertised the general argument: that concerning traits of character. Our ordinary conception of the psychological realm allows for a certain indeterminacy. For example, you may say to someone "Either you are coming with us or you aren't" in order to induce them to decide what they're doing. In one sense, as referring to future events, it is tautologous, but in another, as referring to the person's intentions, it will probably be false - if they haven't then 'made up their mind'. More importantly for our present concern, however, is the fact that what counts, ultimately, in whether some psychological predicate can be truly ascribed to an individual, is their manifesting the appropriate behaviour in the appropriate circumstances. It seems undeniably connected with the ordinary meaning of "Jones is good at learning languages", for instance, that its truth depends on whether "If Jones were to attempt to learn another language, he would quickly succeed" or its opposite is true (cp.[31],p.91). But, even if we don't intuitively conceive of "good at learning languages" as somehow being as real as solubility, say, we might as well conceive it so unless we are prepared to challenge the following general realist picture of the meaning of predicates. Predicates - so eg. "good at learning languages" - delimit particular classes of entities so that, ignoring vagueness, either an entity, such as Jones, falls within the class, or he falls outside it (cp. §2.1). We think of the 'lines of projection'

(the correspondence relation) as reaching out from state-ment to state of affairs (= truth), or not, as the case may be. Our access to that state of affairs may be indirect; we can't just look at Jones and tell whether he has this property, as we can with "fat", say. Nevertheless God, if he existed, could: trivially, saying that the state of affairs determinately exists or does not is to say that a being who could see into that sector of reality would see which - just because its there (or isn't, as the case may be).

In so adhering to bivalence, the realist risks the charge of ignoring what "good at learning languages" really means - the meaning it has for us. For if according to the ordinary meaning of the phrase the ability is, as we might say, 'purely dispositional' - if Jones's being a good language learner consists in, ie. is nothing more than, the fact stated above by the conditional - then we will have no guarantee that Jones determinately either has or lacks the property. For we will have no guarantee that one of the conditionals is true. The point can be accentuated by imagining Jones now to be dead and such that he never encountered anything but his mother tongue. The first assumption means that the versions of 4) and 5) we are now concerned with are past tensed, eg:

- (9) Jones was good at learning languages \equiv If Jones had attempted to learn a language, he would have quickly succeeded.

The second ensures that the conditionals are contrary-to-fact. This assumption does not necessarily mean that neither of them is true. Jones might have displayed qualities - a certain kind of intelligence, say - which is

significantly correlated with the linguistic skill. But suppose we look for such evidence, something to support one or other of the conditionals, and are unable to find any. Two different possible states of the world would account for this epistemic situation. One is that there is evidence around somewhere, but we have been unlucky in not uncovering it. In such a situation Jones would have been a good language learner (or not, depending on the evidence), even though that ability was never called into play, and even though we never discovered it. There is a fact which we could know, and which we would count - via the conditionals - as rendering true one disjunct of:

- (10) Either Jones was good at learning languages, or he was not.

In the other case there is no such fact - there is nothing which, if we knew it, would be counted as grounds for asserting one or other of the conditionals. What are we to say about this case? It is tempting to say that this time (10) would not be true. Early in "Truth" Dummett questions the wisdom of this response of assuming that there is such a thing as a third truth value (for atomic sentences). The notions of truth and falsity are rooted in the more primitive ones of the correctness/incorrectness of assertions, of designated versus undesignated values - and these exhaust all the possibilities. Any putative counterexample that would be recognised by us would really count as a second kind of truth or a second kind of falsity, depending on the case. But this move does not apply here, precisely because this is not a recognisable counterexample for us - we can't tell which of the two stories correctly accounts for our current epistemic predicament. For us,

10) may be true, and may be false (cp.[33],p.149). The realist who insists that 10) must be true must either maintain that the question of Jones's language learning ability doesn't simply reduce to the subjunctive conditional, contravening the meaning we have so far agreed with the anti-realist that it has, or, by insisting on 2), allows that a counterfactual may be barely true. In the former case, he allows for truth in virtue of something which apparently we don't take as relevant; in the latter, for truth even though there is nothing we would count as relevant - both accounts of truth conditions which would be at odds with the meaning we have conferred on it.

Reverting back to the present tense case, ie. when Jones is still alive, rejection of this realism amounts to not regarding him as necessarily good at learning languages or not, prior to anyone's testing him for it. Fulfilling φ and finding that ψ , say, holds good, does not necessarily mean that the conditional $\varphi \Box \rightarrow \psi$ was true all along.

Jones may have been a good language learner, but alternatively it may be that we are not discovering an antecedently existing ability, rather we have created the opportunity for it to come into being.

This last thought is reminiscent of the conception Dummett suggested as a general anti-realist picture of reality: "Our investigations bring into existence what was not there before, but what they bring into existence is not of our own making" ([33],p.18). This is intended as a middle ground between realism and idealism, and as such is puzzling: it attempts to reconcile two notions, discovery and invention, which we are accustomed to contrast. But it suggests a notion of truth as what has been verified,

as proof, rather than what is possible to verify, as provability. On what Dummett takes to be the idealist position, truth does require the existence of proof - to be is to be perceived, rather than perceivable - but the notion of truth is not at all objective. The world is our creation, standing to us as we conceive the work of fiction as standing to its author. Not only is there no notion of truth or falsity independent of our deciding which, but such decisions are also totally unconstrained - we are free to choose. The anti-realist picture drops this last clause; proof is more like something we intuitively appreciate as proof, objectively constrained in some sense, but still the answer it provides cannot be considered as the discovery of some state existing antecedently to its execution.

According to the preceding discussion this metaphysical view cannot be the whole anti-realist picture, though it is perhaps what is distinctive of it. For by the argument so far, if the anti-realist's account of the meaning of disposition ascriptions is correct, we cannot say that the world does spring into being wherever we investigate it, only that it may do. As McDowell says, "a claim to know, in the absence of evidence, that a sentence is neither true nor false [cp. that Jones is neither good at learning languages nor not good] should be no less suspect than a claim to know, in the absence of evidence, that a sentence is either true or false" ([67], p.49). Of course, we have yet to be shown the precise form of clause a systematic verificationist theory of meaning would offer for a dispositional predicate such as brave - or, indeed, for any other kind of expression (with the

possible exception of the logical constants). Clauses which took a form roughly along the lines of 4) and 5) (perhaps characterising ML predicates "Verified" and "Falsified" on the LHS) look like promising target equivalences, because a conditional $\varphi \Box \rightarrow \psi$ spells out conditions for the correct assertion of χ - what one would have to do to get in a position to verify it, and what one should then observe. But as Dummett has emphasised (eg. at [33], p.360) in general the statements which describe the conditions under which we would recognise a given problematic statement as true could be infinitely many. A theory of meaning could surely not employ an infinite disjunction of statements (the truth of any one of which being sufficient for the truth of the statement χ in question.) The lack of any positive proposals concerning meaning specifying clauses remains an obvious lacuna in Dummett's account. Nevertheless, the realist has enough in the thesis built around the claim that equivalences like 7) hold in virtue of meaning of the predicate in question to launch a counter-offensive.

Let me first try to scotch one form of realist rebuttal. The realist notices that whereas χ makes no mention of observers, the RHS's of 4) and 5), the conditionals, do. As Mellor says, "To call a piece of cardboard 'triangular' seems clearly to say how it is in itself, not merely how it is disposed to behave in this or that situation"² - including how it would appear to people. Given this discrepancy, this realist says, there should be no question of the two sides getting anywhere near sameness of meaning. But this faintly Kripkean style of argument is risky when applied to qualities other than primary ones

(perhaps we could take it as a test for primality). Kripke uses something like it to argue that yellowness is not a dispositional property ([59], fnn.66 and 71 - though "related to" one - this much is consistent with what was said above for redness). For Kripke, yellow refers to "that external physical property of the object" which causes in us the visual sensation of yellowness. The sensation we have of yellowness only serves to fix the reference of yellow, not to provide its meaning. But the assumption that there is a unique physical property of the objects is crucial, unargued for and in fact highly dubious, given that physicists make a distinction between surface and pigment colour. (Think also of the blueness of, say, blue paint as against the blueness of the sky.) Kripke also argues that since, if we had different neurology for example, yellow objects would have produced in us a different sensation, yellow cannot be synonymous with "our visual impression of yellowness". But as already argued in section 2.4, this form of argument is a non sequitur. The fact that a descriptive phrase expresses a contingent property of (in this case) a property does not entail that the description cannot be synonymous with the term for that property. Kripke does not show us that we can talk of yellow objects in a world where there are observers with a different neurology etc. other than, ultimately, by reference to how those objects would look to us. This is not to deny that yellowness is "right out there", as he puts it, if this means that whether or not something's being yellow is an objective matter (barring vagueness). The objectivity of colour ascriptions is consistent with the thought that things aren't yellow, red and so on

independently of there existing suitably equipped observers.

The idea that the equivalences 4) and 5) are not beyond dispute, as Dummett would have us believe, is nevertheless worth pursuing. To find a better route to a realist rejection of the anti-realist's account of dispositions, it is worth reviewing some of the empiricist tradition on the topic. Dispositions introduce a multitude of "threats and promises", in Goodman's phrase, and for this reason the empiricist suspects them. A dispositional property should be truly ascribable to an object even when it is not directly manifesting the relevant trait. One has to turn to event predicates to find the non-dispositional. Nevertheless, it may be said, what ultimately counts with solubility, for instance, is dissolvings on immersions. That is, its meaning seems to revolve around correlations of observable events. But as Carnap noted, it cannot be given an explicit definition in extensional terms: the most plausible truth function (\rightarrow) of these most plausible candidates (v is immersed, v dissolves) suffers from the obvious defect, the solubility of the non-immersed. One must renounce either extensionality or the aim of an explicit, eliminative definition, if not both. We have been considering the former option, whereas Carnap stayed extensional, and proposed his famous reduction sentences in place of an explicit definition. In general, these come in pairs, one stating a sufficient condition for the disposition ascription X , the other a necessary condition:

(11) $\varphi_1 \rightarrow (\psi_1 \rightarrow X)$

(12) $\varphi_2 \rightarrow (\psi_2 \rightarrow \neg X)$.

Of particular interest in Carnap's account is his

attempt to do justice to certain diachronic features of language use: the introduction of a new dispositional predicate into a language and a consequent evolution in its meaning, "in accordance with the intentions of the scientist". He recognises that if we make such an introduction, by establishing a pair instantiating 11) and 12), 6) cannot be guaranteed. For "the meaning of $[X]$ is not established completely, but only for cases in which the test condition is fulfilled. In other cases ... neither the predicate nor its negation can be attributed" ([10], p.8). With reference to the class of immersed things (φ_1, φ_2) we can apply the predicate soluble bivalently/classically, for here dissolved (ψ_1) and didn't dissolve (ψ_2) exhaust all the possibilities³ but, it is supposed, we have as yet no way of applying it to things outside that class. We cannot say of a match that has been destroyed by burning, that it was either soluble or insoluble. It is worth pausing to contrast Carnap's burnt out match with the deceased Jones. The reason this failure of classical principles does not have, for Carnap, any "profound metaphysical consequences" is his acceptance, at least as far as these freshly introduced dispositional terms are concerned, of the old-style verificationist equation of meaning with the method of verification - no method, no meaning. In the region of indeterminateness there can be no question of evidence turning up, by chance perhaps, because as Carnap explicitly states, with regard to that region we have yet to confer a meaning on the predicate. With respect to all meaningful uses of it, of course, classical principles still apply.

One might think that Carnap could restore the general validity of 8) by extending the falsity conditions

of X to include the intermediate region. But it would be an arbitrary stipulation to do so at this stage. It suggests as misconceived the scientist's aim to leave it open for us to discover more about the dispositional property in question. It would be to deny that there is any natural way to increase our understanding of X , a way of diminishing the region of indeterminateness which is sensitive to the way of the world. Instead, the scientist is free to define what he means by a term. A theory would be divided between those laws constituting a definition of what it is to be an f , or have property Q , and therefore analytic, and some synthetic principles concerning where f 's happen to show up, which things happen to possess Q , and other contingent features. At worst a scientist could only be mistaken about the latter - he could not produce a false account of what f 's are (their constitution), only a coherently articulated concept to which nothing answers in reality (cp.[86],pp.22ff.). In which case he must scrap that theory - the terms simply fail to refer - and try to mould a new definition. There is no continuity to this process, since each theory constitutes a logically distinct definition ([10],p.12). "Testability and Meaning" represents a step away from such a 'definition account' of theories. It depicts the scientist as slowly closing the gap in the meaning of a predicate, by discovering laws that convey new criteria for its application, these being laid down as additional reduction pairs. Thus, supposing we know what soluble means only for things that have been immersed, "we may perhaps add the law stating that two bodies of the same substance are either both soluble or both not soluble"([10],p.8). Since we know that wooden

things are normally insoluble, this addition would non-arbitrarily extend the falsity conditions of X to include the burnt out match.

Two brief comments are perhaps in order here. The first is that Carnap is not putting the 'being of the same kind' idea to the same use as Quine puts it to (cp.[90], p.224). Quine defines soluble as "being of the same molecular structure as something which has dissolved." But it is not inconceivable that there is a kind of thing whose members are soluble no one of which has been immersed⁴. Carnap is not vulnerable to this accusation, since he recognises this possibility and does not assume that his principle would close the gap completely. Perhaps Quine's account suffices for something like the assertion conditions of solubility ascriptions, since if we are to know whether the members of a kind are soluble, the crucial test is still the dissolving of some specimens under suitable conditions. But this brings us to the second point: many different kinds of things exhibit solubility - what counts, it is still admitted, is the dissolving on immersion. So isn't this a good case for the 'definition theorist'? In the end we may find that all soluble substances share a common kind of atomic bonding, but it is not guaranteed at the outset that such a physical explanation of the property of solubility will be forthcoming. (Moreover, for the definition theorist, saying that things belong to the same kind only looks like a solution. All it does is pass the buck, "for just when are two things of the same kind?" ([44],p.44). I am assuming that Kripke and Putnam have shown enough about reference to show that we do have some right to pass the buck in this way.)

On the "Testability and Meaning" picture, science begins to look like an activity the purpose of which we (being realists at heart) can sympathise with: a progressive attempt to come to understand better the nature of f's. But the picture is not beyond criticism. The idea, which Carnap endorses, that the scientist, in adding these terms afresh with no antecedent meaning creates a new language from an old one itself defines "language" in a way which falsifies the flexibility of natural languages. More importantly, the model is linear, and definition is still considered to be the limit of the cumulative process, ie, when the term's application is settled for all cases (cp. [10], p.13, [44], pp.47-48). We can thus regard the sets of reduction pairs accumulated on the way as partial definitions. Difficulties in this as an account of the ordinary meaning of disposition terms have been brought out by criticisms made by J.C. D'Alessio ([12]). He points out that with a predicate such as soluble, although we can state a principle sufficient for its attribution (dissolving), there is nothing, corresponding to 12), sufficient for its denial. It may always be that a lump's failure to dissolve is not due to its insolubility but rather to the absence of some relevant condition, such as the temperature's not being high enough. Goodman made basically the same point in connection with the subjunctive conditional account of dispositions, that we are "forced back to some ... fainthearted counterfactual" beginning "If all conditions had been propitious and ..." ([44], p.39). (We could get the same effect using a weaker form of conditional connection, eg. with should rather than would.) D'Alessio uses the point to challenge the linearity of

Carnap's model. We need to replace 12), not merely to supplement it, by a reduction sentence which demands the presence of the propitious circumstance if insolubility is to be correctly ascribed on the basis of a failure to dissolve. But since the possibility remains that we shall discover that other conditions are required for a soluble thing to dissolve, it is always possible that the replacement will in turn need replacing. Nor is this just an affliction of the necessary-condition-stating reduction sentences. For D'Alessio goes on to point out that something may not be fragile just because it breaks. Here it is the correlate of 11) that is at risk - pressure, temperature, and the kind of thing which the object is are all relevant conditions. With other disposition terms, eg. magnetic, both kinds of reduction sentence may be subject to correction.

This shows that a simplistic instrumentalist account will not suffice for ordinary dispositional predicates. Solubility does not entail dissolving in any parcel of water, as Ryle once suggested. If a lump of what we take to be ordinary salt fails to dissolve in what we take to be a normal glass of water, the natural response is not "Here is something insoluble, and it happens to be salt", but to wonder whether the lump is salt, or the water is unsaturated, etc. Soluble is a flexible term whose use can be made more precise by explicit relativisation to a particular solvent and, if needed, a particular temperature. What kind of solubility was D'Alessio talking about? Obviously there is solubility relative to the conditions at the test immersion, which quite trivially a thing will fail to possess if it fails to dissolve then and there.

D'Alessio's point must be that the test conditions may diverge in unsuspected ways from what is taken to be the norm. If that is so then he should also have pointed out that they can be too lenient, for instance if the temperature of the water was high enough to dissolve a substance insoluble at room temperature, where what was at stake were the latter conditions.

The important fact - pace 'definition accounts' - is that the evidence that something belongs to a kind whose members are soluble under the relevant conditions is taken to outweigh a particular failure to dissolve - we take it as evidence that the conditions aren't right, rather than that the thing is insoluble. In a complex world, causal mechanisms may break down, or interact with one another in such a way that they are prevented from displaying themselves as they normally do. The failure of some observable feature to occur may be due to the obstruction rather than the absence of the property. Of course, Carnap's law that things of the same substance are alike as regards solubility is not sacrosanct, and it might always be that we have hit a counterexample. Let us consider a case where a physical disposition is not shared by all members of one kind. Taking glass as our substance: some panes are brittle, but some are strengthened. As Hempel says: "The report that the pane was struck by a stone explains its being broken only in conjunction with the additional information that the pane was brittle" ([51], p.458) (assume the stone was not hurled with extreme vigour). It may be that we only know it was brittle in virtue of its so reacting - our easiest access to the knowledge that it is that type of glass coming from putting it to the test - so

in a sense Hempel's explanation is trivial, but it is still true that it was in virtue of its being brittle glass that it did break.

Ordinary talk of the world is suffused with the idea of sufficient reason. An event, such as a dissolving, is not something which just happens. Soluble does not simply gather together actual and potential observable events as brute facts - herein lies the failure of naive instrumentalism (as, to repeat, a theory of ordinary meanings).

Quine puts it thus:

men talked equally easily of solubility before ... explanations [in terms of atomic structure] were at hand; but only because they already believed there was a hidden trait of some sort, structural or otherwise, that inhered in the substance and accounted for its dissolving on immersion. ([90], p.223)

To say that opium puts people to sleep because of its dormitive virtue is vacuous on two counts. The first is the obvious one: at the time it was put forward, it pretended to be what it is not: an explanation - since no independent characterisation of the 'virtue' was available. Secondly, we expected a 'virtue' all along - some feature which opium possesses which gives it that effect on people. It isn't just a happy coincidence, it is counterfactual supporting too: if you had taken that piece, it would have put you out.

What this shows, against the views of Carnap and, in a different way, of Dummett, is how the classical reasoning some group of people adopted in talking about a disposition could be unjustified in the sense that they did not then possess the means always to determine whether or not it was possessed by an object, but for their faith in its existence, realistically construed, to be borne out by

subsequent discoveries, discoveries which would lead them to a fuller conception of its nature. Their insistence on L.E.M. even when they are not capable of proving one or other disjunct requires a picture of the disposition nevertheless either present or not, rendering true one or other disjunct - a picture which they may not then be able to substantiate but which may be correct, if there is only one way of completing the meaning of the predicate, ie. if there objectively is a determinate property which they so far only incoherently conceive. There would then be no route in general from an account of the meaning conferred on a disposition predicate at some stage to the possible lack of determinateness of the corresponding aspect of reality. That not all dispositions conform to the anti-realist metaphysical picture is a conclusion with which Dummett concurs ([31], p.93). But what, finally, of psychological dispositions, for which the anti-realist account was granted to have some credibility?

Quine resolutely upholds the realist faith in connection with dispositions "such as intelligence, whose physical workings we can scarcely conjecture; the dispositional characterisation is all we have to go on" - in this case, and oversimplified, the ability to learn quickly. But he still means by it "some attribute of the body, despite our ignorance concerning it; some durable physical state, perhaps a highly disjunctive one" ([92], p.157). Intuitively, the prospects for realist faith in the existence of a base for the disposition being rewarded seem less in this domain than they do with a physical property like solubility. For one thing, the idea of being of the same substance/kind, so helpful in the physical case, has

less appeal in the mental. Psychological traits exhibit a continuous gradation in the extent of their possession. So if physiological factors are at work, one would expect there to be a combination of several or many of them - as Quine said, intelligence might be a "highly disjunctive" state. What this really means is that even though the physiological realist may agree with the correlation of the disposition with the related behaviour-specifying conditionals 4) and 5), he may not want to commit himself to 6) (and therefore to 2), etc.), on the grounds that these predicates do not have cut and dried conditions of application. If there is an element of vagueness in "intelligent", one may decline to accept "Either Jones is intelligent or he isn't" on the grounds Jones may be a borderline case. Recognising that intelligence covers a spectrum of cases would thus seem to defuse the dispute, as set up by Dummett.

I am inclined to think that neither a physiological realist account nor a purely behavioural account of the meaning of these predicates will tell the whole truth. Both types of account are abstractions from our linguistic practice - practice which exhibits tendencies in both directions. On the one hand, someone may think he 'really' possesses a certain quality - wit, say - even though he fails to manifest the appropriate behaviour given the opportunity. He fosters the idea of a trait enduring through its failure to be manifested. On the other hand, those who disagree with him may take it that in part his failure to produce the goods is evidence against the presence of an underlying disposition, but in part also just constitutes not being witty. Might it not be that we are prepared to

bend our conception of what a mental trait consists in according to our particular interests?⁵ The members of a linguistic community are not taught words uniformly - which words are basic, and which defined, will vary from person to person, as will the precise wording of the definitions, where given. Dummett's claim that a predicate like "good at learning languages" is "introduced" via a subjunctive conditional, though plausible, is of unclear status as a hypothesis about language. As a piece of armchair linguistic psychology, it seems to me more plausible to say that what we are initially exposed to (with a predicate which is unstructured and which is not explicitly defined for us), are positive and negative exemplars. How precisely we extrapolate from this is an empirical matter (as indeed is the question of whether such training does occur) which should leave us with some freedom in how we understand the word.

3: POSSIBLE WORLDS AND ANTI-REALIST MODALITY

In this section I want to consider how the difference that should arise between the realist and anti-realist in their accounts of necessity and possibility will affect the possible worlds interpretation of these notions. We have already (section 1.5) looked at one way in which possible worlds talk can be regarded as an extension of our more everyday operator language. If, being a refinement of ordinary modal talk, the possible worlds rendition of it is to this extent unexceptionable, a suitably modified version of it should be acceptable to the anti-realist. But first let us consider the kinds of demands anti-realism makes on the modal notions themselves.

In his exegesis of Frege, Dummett phrases the discussion of the interpretation of necessity in terms of the notion of analyticity. For the anti-realist, any notion of necessity must relate, like the notion of sense, to the means we have of recognising sentences as true. So he favours epistemic analyticity - roughly, truth in virtue of the meanings to be found in our language, anti-realistically construed. By contrast, realist necessity will typically be non-epistemic like realist sense, relating to the kind of thing in virtue of which a sentence will be true, whether or not we could know it ([29], p.117). The realist should favour ontic analyticity - epistemic analyticity, we could say, relative to God's language (= epistemic analyticity relative to our language, realistically construed.) The idea of God's language is that it is one which would be spoken by an omniscient being, one who sees things as they really are, and whose language

therefore describes things as they really are. It is deliberately intended to enshrine several (potential) myths.

According to Dummett, Frege retained the Kantian trichotomy of analytic, synthetic a priori and synthetic a posteriori. This is somewhat misleading, since Frege adopted a wider notion of analyticity than Kant: a sentence is analytic if it can be proved from definitions and logical laws ([29], pp. 502, 632 and [41], section 3). But even this Dummett regards as too restrictive, and at one point he suggests that a sentence is analytic if the procedure for determining its truth value can only yield one result ([29], pp. 655, 636): our discovering it to be true, presumably. Of course, any true sentence, so described, is such that we could only discover it to have the truth value true. But the point is that an analytic sentence has that status in virtue of the sense it possesses, not the way of the world.

Since our recognition of the truth of an analytic sentence may not be immediate, such a sentence may be informative. The informative/trivial distinction cuts right across Frege's analytic/synthetic distinction. There are analytic trivialities, " $1 = 1$ " being a good example. Fregean sense, according to Dummett, is transparent to users of the language: if two words have the same sense, and someone understands them, he must know that they share that sense. So a genuine expression of synonymy such as " $1 = 1$ " must be trivial, i.e. recognition of its truth will follow merely from a grasp on its meaning. (This is just the contrapositive of: because "Der Morgenstern ist der Abendstern" is informative, the two Eigennamen, as Frege

called them, are not synonyms.) By wanting to prove that all true arithmetical equalities are analytic Frege had, in analytic equivalence, a notion which obviously goes beyond synonymy: most of these equalities are informative. Now consider "Either the sun is shining now or it isn't". We could verify this by looking out of the window. That kind of verification would be on a par with that for, say, "Either the sun is shining now or it is snowing" - it follows directly the composition of the sentence out of its disjuncts. Its truth would have been ascertained, but not its analyticity. It is unlikely that an ordinary speaker would consult the weather in this simple case, but it becomes more probable the more complex the analytic sentence is that one might directly verify it without realising it was analytic ([29], p.635, and cp. p.120).

As Kripke has pointed out, people can learn arithmetical truths a posteriori, by consulting a computer or a mathematician ([59], pp.261,765). The computer may supply us with answers that transcend what it is humanly possible to calculate - ie. truths which are in practice impossible for us to ascertain a priori. This example of a necessary truth being verified a posteriori does not impugn the idea that the most direct and conclusive route to its verification is a priori - it was a priori for the computer (or mathematician). That idea is, rather, brought in question by the previous, meteorological example. Clearly, the notion of a priori requires a modality in its characterisation (as had analyticity above): a truth is a priori if it can be known - whether directly or indirectly - without recourse to experience of the world. (This modality can perhaps be eliminated in favour of a quantification over

methods of verification.) Even though we may be tempted to conceive of apriority as a species of necessity, thus construed it is a problematic rather than apodeictic modality. And even when we acknowledge the confusion Kripke mentioned, of substituting "must" for "can" in "can be known independently of experience", the temptation remains to a certain extent.

One reason for this is that it is natural to think that apriority is at least coextensive with a notion like epistemic analyticity. For how else could one derive knowledge independently of experience if not ultimately from a grasp of the meanings of the words? - and conversely, if a statement is epistemically analytic, it should at least be possible to know it independently of experience. To be weighed against this is the well-known dissociation of apriority from metaphysical necessity argued for by Kripke. If we think of the concepts in a series ranging from the most to the least epistemic: apriority - epistemic analyticity - ontic analyticity - metaphysical necessity, then if a connection fails to hold in either direction between the outer pair, some links must fail in between.

Let us consider some cases. Concerning Kripke's rejection of: if it is metaphysically necessary that φ , then it is a priori that φ , it is worth observing that one case which might spring to mind, "Hesperus is Phosphorus", is not in fact so obviously a counterexample. Suppose that someone understands of the pair of names only "Hesperus". One way of imparting to him an understanding of the name "Phosphorus" is by telling him "Phosphorus is the planet Hesperus" or "'Phosphorus' is another name for Hesperus". He will then acquire the knowledge that Hesperus is

Phosphorus, doing so independently of experience (other than that involved in acquisition of a grasp on the senses of the expressions), by deriving it from his understanding of the words. An a priori truth is, to repeat, one which can be known independently of experience. Perhaps that characterisation can be qualified (to exploit the fact that, as the names are supposed to be have been introduced, it required observation to know they stand for the same object - cp.[60],p.73). No matter - one could pick other examples from Kripke's work not subject to this kind of consideration¹. Moreover, Dummett is prepared to concede that there is no entailment from something's being ontically necessary to its being a priori - provided that a realist interpretation of our language can be made out. In fact, his position here seems to allow the possibility, cultivated in the last section and to be discussed more fully in the next, of there being more to a metaphysical issue than can be extracted from a theory of meaning for our language. The main example starts from the plausible assumption that were God to exist, his existence would be necessary in some metaphysical sense. Aquinas believed that this could not be reduced to any linguistic species of necessity stemming from our reflecting on the meanings of our words (cp. [29], p.118). Against this some infidel apparently argued that if a statement is to be necessary, it must be analytic, and if that is to be so, since the statement "God exists" comes from our language, it must be analytic for us. Assuming with Aquinas that this consequence is false, it follows that God's existence is, at best, contingent. But this argument is not compelling for one who takes the realist point that there can be more to a metaphysical issue than

meets the semantical eye (at least as trained on our language). If we have to think of necessity as some form of analyticity, this means that there can be a notion of ontic analyticity distinct from epistemic analyticity. The infidel has offered us a conditional: if a statement is not analytic in one language (eg. ours) then it (ie. anything we ought to count as a correct translation of it) will not be analytic in any other language (eg. God's). What this ignores is the possibility of its conditions of truth being fulfilled transcendently by our standards, and hence, if those conditions cannot but be fulfilled, that the statement is analytic in a language for which a realist interpretation is permissible (by the anti-realist's lights - ie. coincides with the account he would offer).

Another example to the same effect would be provided on the supposition of the correctness of the platonist view of a transcendent mathematical reality ([29], pp. 117-18, 120). Since arithmetical statements are, if true, necessarily so, undecidable statements which are nevertheless true in virtue of that platonic reality would count as ontically but not epistemically analytic. The obvious example here would be an arithmetical statement involving unbounded quantification over numbers. As regards finite cases, the universal quantifier will be associated with the same direct and indirect means of verification in our language and God's, but with infinities only the indirect means will be shared. The statement will fail to be analytic for us because we are not in possession of a procedure that can only terminate in our recognising it as true - the direct means of verification, surveillance of the totality.

Surely, a truth can be known independently of

experience only if no observation we could make could count against it. That is, it's not that an a priori truth can't be verified a posteriori, it's that it can't be falsified a posteriori. One might then think that, if in no experienceable situation it will come out false, in every such situation it will be true. Apriority would be a special case of epistemic necessity - truth in all situations compatible with what we know, once all the a posteriori knowledge is subtracted. But this conflation is unsound. Falsity in an epistemic situation should be recognisable as such, truth likewise; the fact that we could not experience a falsification does not mean that we cannot but verify it. We have just seen examples which for the anti-realist emphasise that there are statements which we cannot falsify a posteriori and which, because undecidable, are not verifiable at all, hence not a priori. The possibility of realism means we can add: unfalsifiable because undecidably true, ie. metaphysically necessary.

This still leaves the converse connection intact, from apriority to metaphysical necessity. It is the challenge to this which Dummett finds most objectionable. A source of such antipathy is diagnosed by Kripke: if a statement belongs to a class whose members are, if true, contingently so, we should not be able to rule out the possibility of a falsifying state of affairs occurring without recourse to checking the world ([59], p.263). What this consideration ignores is that one can pick out a contingent state of affairs a priori by a judicious choice of wording (for more on this, see [39]). We all know ab initio that wherever we are, it will be correct to call that place "here", but this does not mean that we would

regard "It is necessary that I am here" as true. Consider the query Dummett raises concerning the individuation of the facts involved - about what fact it is which is supposed to be both known a priori and contingently true ([29] pp.122-23). It is plausible to say that, for example, an utterance of "I am here" does not state the same fact as α is at β , where α and β are names of the relevant person and locality. The two statements clearly have different properties - the latter does not embed truthfully within the context "It is a priori that". But this is not enough to overturn Kripke's challenge. What we are primarily concerned with is whether "I am here", although embedding truthfully within "It is a priori that", does not do so with "It is necessary that". It may indeed be that "It is analytic that I am here" is true - a case of modality de dicto, since one infers the truth of "I am here" from an understanding of the words. But this can only encourage us to think that the contingency of "I am here" is not a matter of the denial of a language-relative notion of necessity (which is perhaps even more apparent with an example like "I exist"). In the previous case, we could go along with making necessity a kind of analyticity, by appealing to the idea of a possibly transcendent language. The idea is inapplicable in this case. The fact that we may read "It is necessary that I am here" as false shows that we are prepared to treat it as an instance of de re modality, of the contingency of my being here.

So far, then, we have seen no reason to think that a suitably epistemologised notion of possible world (cp. the "non-realist" worlds suggested by Kamp in [54]) would not provide a better account, for the anti-realist, of the

notion of necessity than an elucidation of it in terms of some overtly epistemic notion like apriority. At one point Dummett does suggest that such a notion would provide a properly cognitive representation of sense ([29], pp. 134, 135). Before considering what the anti-realist would regard as the unacceptable face of possible worlds, traditionally conceived, it will be worth reviewing their better side. Suppose that someone wanted to defend in detail the claim that Nixon could have been a used car salesman. The natural way to do this would be to find out the facts of Nixon's past life, and then modify in thought those circumstances which inclined him towards a political rather than automative vocation. What our speculator does, in effect, is to stick the modality out front (eg. "Things could have turned out like this:...") and then to describe in the indicative and in the detail he requires the way the world would have been. He could go on to spell out other possible denouements and make comparisons between them (which requires some means of cross-reference). All this provides more detail than our ordinary interest in modal matters requires, though we may suppose that military strategists, concerned as they are with what could have or could happen, avail themselves of this mode of speech (albeit with possible 'scenarios' rather than 'worlds').

One must keep a sense of proportion. The technique of exploring the modal realm by rewording things in terms of possible worlds gives rise to a Scheinproblem, that of trans-world identification. We imagined a world in which Nixon was a used car salesman. But as it happens Nixon does not enjoy this property - so how, by Leibniz's law, can the character in the counterfactual situation be

identical to Nixon? Something must be amiss in this question, for Leibniz's Law is not a modal principle; it tells us that a thing is what it is, not in the sense that it couldn't be any other way than it in fact is (which is where it leads, on the reasoning behind the question), but only in the sense that difference of properties suffices for the actual numerical distinctness of any two individuals. Our problem was simply whether "v could have been a used car salesman" is true or false of Nixon. This is the basic issue - whether an entity a has a property φ in another world is nothing over and above the question of whether $\Diamond\varphi a$ is true. If, by rewriting the problem in terms of possible worlds we find we are unable to describe a world in which Nixon has the desired non-modal property, this will count against ascribing the modalised property to him. It is not, then, that the method of possible worlds churns out negative answers to every question of the form "Is $\Diamond\varphi a$ true?" (where in fact $\neg\varphi a$) by automatic application of Leibniz's Law.

As Kripke says, we want just to stipulate that we're talking about Nixon in another situation. This doesn't mean that we can simply stipulate that an individual is a particular φ thing in an imagined situation, whatever the value of φ , for that would amount to the converse of the previous error, making $\Diamond\varphi a$ always true. Perhaps Kripke's point is this. When making ordinary, non-modal assertions if what we say is not true, it's not that we have made a true assertion about another world, we have made a false claim about the actual state of things. But when speculating counterfactually, there is no question of our making a false claim about the world we want in that manner.

There is no independent way to refer to a non-actual world except through the description we make of it². So we are free to choose the individuals we want to talk about, and describe them as having the properties we want to imagine them as having - provided, of course, that those descriptions stay within the bounds of genuine possibility.

At least this much, then, can be said in favour of the innocuousness of possible worlds. Yet an anti-realist would find them objectionable on three counts, I think, as standardly used in the semantics of modality. Firstly, all the realist assumptions about this world are carried over to possible worlds, on the traditional conception. For example, Robert Adams explicitly characterises them as determinate ([1]), meaning by this that a) every world is such that exactly one proposition out of each contradictory pair is true with respect to it, and b) each world, if temporally ordered, is a complete world history. This 'maximal consistency' claim is shared by Plantinga, though formulated in terms of states of affairs³. The point here is simply that possible worlds will be vulnerable - susceptible to non-determinacy - at whatever points the anti-realist establishes a case independently of modality. And as such, they will be unsuited to be used in any theory of meaning which the anti-realist would put his name to. This is brought out nicely by a ploy of Newton-Smith's ([74], pp.68, 233-34, and cp. §3.4 below). Suppose we have a case of underdetermination of theory by (all possible) data; that is, two theories T_1 and T_2 for which, while they apparently make incompatible claims about the nature of the world, there is no possible observation we could make that determines which one we should select. The realist may

happily accept this predicament, as proof that the truth can, and in this case does, transcend our means of ascertaining it. If so, he will be inclined towards depicting it in terms of two definitely distinct worlds, A and B, corresponding to T_1 and T_2 respectively. Our difficulty is purely epistemological: we can't tell which one we're in. The anti-realist, however, will reject this model: we have no justification for thinking we're in a possible world, if this is what they are, ie. determinately constituted conditions for the truth of one or other of T_1 and T_2 , independent of our epistemic abilities. (Here the divergent paths of Quine and Dummett (cp. section 3.1) momentarily reconverge. But there is a difference: for Quine, what is objectionable is not that, as it were, the worlds are determinate, rather that we should regard their identification, through language and theory, A by T_1 , B by T_2 , as determinate. This remains realist in that the epistemological difficulties are our problem.) This is the objection which Dummett has most frequently raised against possible worlds: their unsuitability to represent the cognitive notion which sense is (eg. at [33], pp.421-22). Only as aggregations of recognisable states of affairs could it be thought that they could contribute to such a project.

Secondly, and most importantly in this context, the epistemology of modality is no easy matter. The idea that some truths are necessary is not readily accessible to an empiricist epistemology, for the obvious reason that we do not perceive necessity in the world. As Kant is oft quoted, "Experience teaches us that a thing is so and so, but not that it cannot be otherwise". Even with Kripke's plausible candidates for the category of necessary a

posteriori, we observe only that eg. Hesperus is Phosphorus - that it must be so is established by the a priori argument that if $a = b$, then $\Box(a = b)$ (cp.[59], p.765).

Similarly, it may be argued, we don't observe that its possible that Hesperus is Phosphorus, we just observe it. If our world is but one among many other possible ones, it is in general known to us as such via the mediation of the ab esse ad posse principle, $\varphi \rightarrow \Diamond\varphi$. Not too much should be made to hang on this notion of the directly observable though. Might it not be said that we can just see that its possible for the piano to go through this doorway? Or that its impossible for it to go through that one? The example in the last section of the impossibility of the burnt up match stick dissolving in water illustrated, in its humble way, that a certain kind of possibility is open to empirical investigation. But the important point is that we have no guarantee that we will be able to conclusively determine all questions as to whether or not $\Diamond\varphi$, for instance⁴.

Possible worlds, realistically construed, present to us a picture which conflicts with this. Recall Kripke's remark that realism about possible worlds represents them as things "one looks upon ... as an observer" as if through a telescope ([59], pp.266,267). The image of an independent realm rendering objectively true all matters modal, awaiting discovery by astronomical means, compares with that presented by mathematical platonism (cp.[33], p.229). But the idea of observation here is not one we can take seriously⁵ (especially if some causal link between perceiver and perceived is required. For how could we receive a causal signal from some non-actual world?)

Another point is that taking necessity as an

unbounded quantifier over an infinity of worlds will only compound the ills already accumulated (cp. section 3.1). One might also wonder whether the anti-realist would query the very intelligibility of a possible worlds ML. For example, a conventionalist would query the idea of a once and for all domain of possibilities, the set W -- even if one adopted an intransitive accessibility relation to define the modal concepts of the OL (to appease the conventionalist's presumed dislike of the S4 principle, $\Box\varphi \rightarrow \Box\Box\varphi$), the ML would still embody an absolute perspective on what is possible. But let us focus on the problem raised by Peacocke ([79], pp.480-81) that a homophonic satisfaction-based truth theory for a first-order language with modal operators would serve up the wrong truth conditions for an OL sentence saying that there could be incompatible objects -- objects which could not coexist in any possible state of affairs. For it would require the possibility of a sequence containing a pair of such objects, and such a sequence could not exist in any possible world. For it could only exist in a world in which all its members existed, and by hypothesis no such world exists. The sequence principle 1.2.12 would be an ontological correlate of the invalid principle $\Diamond\varphi \ \& \ \Diamond\psi \rightarrow \Diamond(\varphi \ \& \ \psi)$. What this line of reasoning ignores, however, is that the peaceable coexistence of such possible entities in some set is already presupposed in the assumption that there is a domain of possibilia A over which the variables of quantification range. Once we have such a domain, the idea of functions from \mathbb{N} to its members must be acceptable. The point is that when dealing with such sequences we are not dealing with a perspective which is bound to a particular

world. The problem serves to highlight the difference between a genuinely world-bound view of existence and one which is modestly supramundane. The difference is most vividly illustrated by the assumption that there is a set of entities over which the possible worlds variables range. Clearly, such a set could not exist in/at any one world for, from a world-bound perspective on existence, that would require all its members to exist in that world. Worlds are not the kind of thing which exist in or at a world. But to assume this much of a supramundane perspective - ie. simply to quantify over worlds - is not what is problematic for the anti-realist. Consider the temporal case. As argued (§1.5), the use of quantifiers over times is a natural extension of our ordinary use of temporal operators. What the anti-realist finds objectionable in a realist construal of time is not such quantification per se but the tacit appeal to the powers of a being who transcends time and who views all times on a par (cp. [33], p. 369). That is the kind of ability the realist would have to appeal to if he is to explain what knowledge of the truth conditions of past tense statements consists in, considered as determinately fulfilled or not, irrespective of our inability to discover which. Because this being stands outside of time, the essential point must be that it is impossible to add ordinary temporal indexicals to the language it would speak - there would be no "now" for God. Such an existence is not one we have any serious conception of. The same mistake is made when we picture God's view of the modal universe - the set of all possible worlds - by analogy with a spaceman looking down on different planets. There is no genuine analogy here. A possible world is a

totality of facts. A being for which all possible worlds are on a par, equally accessible, is thus one who stands outside all facts, from whose vantage point it would be impossible to recover the perspective of the actual world.

We can summarise the anti-realist critique of the classical conception of possible worlds in the terms of a triune exorcism. Firstly, a being who can apprehend a wholly determinate totality of facts, as would be invoked by Adams's requirement b) for their determinacy through time, has powers which bear no real resemblance to our own. Nor can we empathise with a being who can directly perceive more than one possible world - even if restricted to effectively recognizable features thereof. And thirdly, the same goes for a being who watches over an infinity of worlds (and a non-denumerable one, at that). In all these respects adoption of a non-realist theory of meaning would lead to a revision of the concept of possible world.

4: SEMANTICS AND METAPHYSICS

In 1959, realism consisted in "the belief that for any statement there must be something in virtue of which either it or its negation is true" ([33],p.14). More recent formulations put it in terms of a statement's being either determinately true or determinately false (a stronger version of bivalence than the barely stated: every statement is either true or false), irrespective of whether we could ascertain which (cp.[33],p.225). A realism is realism restricted to a particular subject matter. Three important notions are operative within (a) realism: bivalence, potential verification transcendence, and the guaranteed existence of the 'something in virtue of which'. It is clear how these ideas are supposed to mesh together: the last one justifies or entails the first two. Dummett has written:

a conception of meaning as determined by truth-conditions is available for any statements which do relate to an independently existing reality, for then we may legitimately assume, of each such statement, that it possesses a determinate truth-value, independently of our knowledge, according as it does or does not agree with the constitution of that external reality which it is about. ([33],p.228)

The kind of doctrine which provokes the anti-realist is thus depicted as, essentially, a specific metaphysical conception of some part of reality finding expression in truth-conditional semantics for the area of language concerned with it. The notion of truth which is employed, as 'agreement', is not itself under attack. Basically, that is the idea of truth as correspondence, as captured by what Dummett terms 'Principle C', (here strengthened to a biconditional):

- (C) A statement is true if and only if there is something in virtue of which it is true.

This notion of truth is but an innocent bystander, expediting the anti-realist's reductio of some doctrine concerning a facet of reality. To construe matters thus is to place Dummett within the great verificationist tradition of showing up certain metaphysical theories as nonsenses, preferring conceptions neither theirs for the giving nor ours for the taking. And an inevitable consequence is that classical truth conditions cannot provide the fuel which will in general power a theory of meaning over the whole linguistic terrain.

Consider the conditional

- (1) If reality is determinate in certain respects, biva-

lence is guaranteed for the relevant class of statements. An anti-realist would accept 1), and utilise it in a modus tollens argument, ie. when he has ascertained that bivalence is not guaranteed for some linguistic expression. The realist is also likely to accept 1), but argue the other way round - we are naturally inclined to take a premise about the constitution of reality and detach the consequent. We think, for example, that because the material world is determinate, any non-vague description of it we make will either hold or not, epistemological considerations being irrelevant. It also appears as an element in Lewis's realism about possible worlds. He has argued that because worlds are non-invented, autonomous entities, there are many facts about them he does not know, nor even that he knows how to set about discovering ([63], p.88). Nevertheless they remain facts: just because worlds are not figments of his imagination, Lewis implies, these facts are determinate.

It is not open to us to influence them.

Consider now the converse implication:

- (2) If bivalence is guaranteed for a class of statements, then reality is determinate in whatever way apt.

The anti-realist need have no quarrel with this, since bivalence does not necessarily entail possible verification transcendence. He grants that it is unobjectionable to attribute knowledge of classical truth conditions to a language user, the obtaining of which will always be recognizable. Thus Wright:

we do not ordinarily think of those aspects of reality which we are able conclusively to determine as any less of the world. All that is being suggested ... is that such aspects constitute the world - or, at least, those of its features to which we can give intelligible expression. ([113], p.226)

(Note the caveat in the final clause.) Nothing untoward is conceded by this (at least by the standards of moderate anti-realism), since the idea, of which the realist would like to avail himself, of a truth condition that can transcend all possible means of recognition will not be, ex hypothesi, one applicable to the usage of decidable statements, such as reports of observation. Finally, if it is part of the intuitive realist attitude to put metaphysical considerations first, the realist can also use 2), when reasons are adduced for rejecting the consequent. Fiction provides a standard example. We might grant that there is a sense of truth for which "Sherlock Holmes was a detective" is true. But given the belief that no actual or possible individual could be identical to Holmes ([59], p.764), we are likely to conclude that there are sentences concerning Holmes which are neither true nor false (cp.[32], p.385). So bivalence cannot be guaranteed for the notion

of truth - if we are prepared to call it that - as applied to fictional discourse in general (in fact one could actually cite counterexamples to the principle).

If the preceding is right, then both the realist and the anti-realist agree on the equivalence, as defined by 1) and 2) together, of the semantical and metaphysical issues. This tallies with Dummett's position in Elements of Intuitionism (p.386): "the metaphysical question, ... what facts obtain [ie. "hard" facts, "those which constitute the substance of reality" as it is in itself [35], p.222] is the very same question as the question which statements we can suppose to possess a determinate truth value." The point is that the notion of fact is strictly correlative with that of true statement - there is no question of either disputant trying to analyse one notion (eg. truth) in terms of the other. But this potential for agreement is obscured by what has emerged as a clear difference in methodology. The traditional realist prefers an approach which assigns a priority to the metaphysical issue. Perhaps what lies behind this methodology is the idea that reality itself existed prior to our coming on the scene and talking about it. It could have turned out that there were no beings with any linguistic abilities anywhere, so it is surely a contingent fact that beings with our language and our epistemological powers happened to show up. This makes it natural to think that, for all we know, there are features of reality which are beyond our ken¹. But this general moral that we should not make reality itself turn on our possibly impoverished conception of it, is quite compatible with the methodological point that its not open to us to decide first what facts determinately obtain. The

realist could anyway concede that the situation is more complex than would be represented by unidirectional moves from reality to meaning. Take the intuitive opposition to the determinacy of the fictional realm discussed above. Trivially, you can't get to a metaphysical thesis about fictional reality without knowing something about the linguistic practice of writing fiction, telling stories, etc. And to appreciate that fully, you may have to contrast it with more normal discourse which, for the contrast to be effective, may in turn require mention of the non-fictional reality which it concerns.

The situation so far would seem to be reflected perfectly by Dummett's characterisation of the Fregean doctrine of naming. He has written:

Frege's use of the ontological term 'object' is strictly correlative to his use of the linguistic term 'proper name': whatever a proper name stands for is an object, and to speak of something as an object is to say that there is, or at least could be, a proper name which stands for it. ([29], p.55)

But we don't first comprehend somehow that numbers are objects. Rather, there are certain syntactic tests by means of which we can identify proper names, such as an expression's permitting existential generalisation, having no plural forms, and so on. If an expression passes the test, then it is a proper name, and its sense and reference will be fixed. Numerals pass the tests, so numbers are objects for Frege. No further questions about the existence of numbers arise - to think that they do is to make the mistake of thinking that objects are objects independently of our conceptualising them as such through the apparatus of reference and predication. It would be, for Dummett, a transgression of the Fregean edict not to inquire after the

Bedeutung of a word in isolation from the kind of linguistic context in which it occurs (cp.[33], pp.40-41).

This more or less neat picture of the way the disagreement between the realist and his opponent is shaping up is destroyed when Dummett gets on to the question of how a thesis about the externality of reality would relate to one concerning its determinateness, ie. a realism in his sense. Let us first fix a piece of terminology. An issue is ontological if it is of the former kind, ie. if it addresses the question of whether some facet of reality is external to us, or is indebted in some way for its existence to such thoughtful beings as ourselves. A metaphysical (or metaphysical/semantical) issue focuses on the kind of question we have been discussing - the hardness/softness of the facts, or equivalently (by 1) and 2)), the bivalency of the notion of truth for the relevant statements. Dummett's contention is that an ontological thesis - if it can be given non-metaphorical content - cannot serve as a premise from which to derive a metaphysical/semantical conclusion. There would then be two senses in which an account of reality would be secondary to the relevant part of the theory of meaning - in the methodological way discussed above, and in this other sense, that we cannot adopt a stance on the ontological issue prior to determining the correct notion of truth. Having done so, one or other view of reality - external or dependant - will inevitably suggest itself as the natural one (cp.[33], p.229).

In order to demonstrate his contention that metaphysics doesn't follow from ontology, Dummett outlines two hybrid doctrines. As they are initially presented in Frege

(pp.507-8), he does not address the question of how we would flesh out the content of the ontological half of each hybrid. For the one position, someone in effect says:

"Numbers exist autonomously, nevertheless quantification over them is to be construed in terms of proof and disproof, ie. intuitionistically." For the other, they say: "I believe that numbers are mental constructions, nevertheless quantification over them obeys classical principles."

Given the latter halves of these two positions, it is natural to wonder if the former is not, as it were, an insincere mouthing of words. For obviously a dispute about objects doesn't arise for them stripped bare of every kind of property they enjoy, and this brings us back to the nature of the facts they participate in. From a prephilosophical point of view, these doubts may not seem entirely convincing, since a thesis about objects has plenty of intuitively comprehensible content. We have firm intuitions on the nature of objects, shaped by a paradigm: that of the spatio-temporal, physical object. This is surely one reason why the strand of formalist thought that denies outright that mathematical statements are about objects is intuitively appealing (cp. section 2.1). (From this perspective, the Fregean conception is deficient. Even though a generic "the lion" or a species name "panthera leo" may pass all the syntactic tests, we would prefer a semantic account which does not construe them as directly referring to a single object - a platonic Lion, presumably.) That thought is more appealing than either the picture of the mathematician as astronomer, or the picture of him as artist. Dummett rejects this aspect of formalism on the grounds that to say that one kind of discourse is not about

anything, or any things, is to say that one totally fails to make statements in that area ([33], p.xxv). He does not argue for this principle. In one sense it is unobjectionable: by reference to the alternative, Fregean notion of object. It is indeed in such essentially semantical terms, rather than by comparison and contrast with the spatio-temporal paradigm, that Dummett develops the ontological halves of the two positions. It is then unclear why a formalist would or should not maintain that the thesis that numbers are objects is, on this Fregean conception, harmless - assuming that numerals encounter no difficulties with the syntactic tests. He would then grant it unobjectionable to talk of our bringing numbers into being - wholesale, by laying down a system of axioms, rather than retail, 'as we go along'. Whether he would still resort to classical proof procedures would then be just the question of the coherence of the second position outlined above.

Let us now try to develop these two kinds of hybrid in more detail, but by continuing the theme of possible worlds. In one case, possible worlds are not considered as existing "way out there", as Kripke puts it, but as dependent on human thought. Suppose this means 'as we go along', ie. that they exist in so far as we think of them - excepting, of course, the real world. Simultaneously, what is necessary and possible is deemed an objective matter - bivalence holds for modal statements, and they obey classical logic. So presumably necessity, being objective, constrains our construction, in thought, of other worlds. As Dummett says, an object of imagination such as a fictional character can only have those properties it is described as having. There is no question here of the choice of

properties being constrained in any way. But, by assumption, there is no such freedom in this case. The problem of the intelligibility of the position is not yet dispelled - to understand what is meant by saying that possible worlds depend, in this sense, on thought for their existence. If talk of other worlds is just another way of talking about what is possible, and if the latter is objectively determined, then what stops us thinking, concerning any particular question of possibility, that the relevant worlds already existed prior to our thinking about them? It should be predetermined whether or not some description will count as a coherent description of another world. So what do we do to bring them into being, since we are not free to stipulate whether or not something is possible? It looks as if what this theorist really believes is that unactualised worlds don't exist at all, like the formalist just discussed, and so that all talk of 'bringing them into being' is, really, metaphorical. We get another angle on this problem by pressing the question: what justifies the theorist's belief in the objectivity of modal discourse? It cannot be slavish adherence to bivalence, since by hypothesis he rejects that for statements of the form "There is a world such that ...". If he thinks that there is something - other than possible worlds - which renders it bivalent, then this other facet of reality must be the prime claimant to be what modal discourse is really about. This possibility is perhaps more evident with modality than mathematics, for in this case someone might well believe in its objectivity without thinking he thereby concedes the determinate existence of other worlds, precisely because he doesn't believe that modal discourse

carries commitment to them. Realism about possible worlds is not the only view that someone insisting on bivalence for modal discourse need appeal to for support. (This is not to say that one first adopts bivalence without any thought of the metaphysics required to justify it, as if the latter would simply be an afterthought.)

A more plausible, less extreme concoction along these lines would attempt to reconcile a non-provability transcendent notion of truth for worlds with the use of classical logic for modal discourse. If talk of our bringing worlds into existence is to make any sense at all in the context of an affirmation of the objectivity of the modal, it must relate to our bringing in the whole universe of worlds at once - a universe which could not exceed what we have put in or could get out. The idea would be that although bivalence is not guaranteed for statements proclaiming the existence of possible worlds which are such and such - and, therefore, for their more ordinary operator counterparts - we would still be entitled to reason classically about modal matters. This would show that the issue which really counts in a dispute over realism, the question of what kind of logic we are entitled to adopt, could not be decided by the prior adoption of a stance on the ontological issue (*à la* Kreisel's famous dictum).

If this position is acceptable, there must be a justification for classical logic alternative to the principle of bivalence. One natural response to this challenge is to say its use is harmless - as harmless as its adoption when vague predicates are being used. Vagueness could set an important precedent, so it is worth considering briefly. Consider an object in the middle of the

continuum of shades between red and orange. Given that "red" is a vague predicate, we cannot regard the statement "That is red" as either determinately true or determinately false. It is clear that this does not have anything to do with the possibility of our observations effecting a change in it - it will be as much red before as after. Nevertheless, one kind of defense for "Either it is red or it isn't" is that, on any way of making the application of the predicate more precise, one or other of the disjuncts would come out true. So what harm can be done in assuming LEM here? (There is no question of an anti-realist raising the possibility of there being nothing which could count as evidence for one of the disjuncts - the object definitely has a colour.)

We get closer to the theory we are after by considering Dummett's specifically Temporal anti-realist T, who distinguishes truth in a possible past world history from truth in all past world histories (ie. that which is still determined as having happened). A past tense statement may not be bivalent in the latter sense of truth, since the traces of the past which remain may not determine it as true or false. Nevertheless, T regards the statement disjoining it with its negation as determinately true in this sense, because he regards each possible past history as complete, and thus bound to verify one or other of the disjuncts. T's defense of classical logic is very similar to the preceding one: although $P\phi$ may not (any longer) be determinately true or false, there is no harm in asserting $P\phi \vee \neg P\phi$, since whatever possible past history was the actual one (cp. whichever sharpening of "red" you choose), it comes out as true. T's concession to anti-realism is in

agreeing that, on the important notion of truth, bivalence is not guaranteed for past tense statements. Thus although he holds that it is indeterminate which history was the actual past history, because he regards each history as a maximal fleshing-out of what we currently know, whichever one it is it is determinately constituted.

What would be a comparable case for modality? One way to think of it would be this: if philosophers cannot conclusively settle such conceptual matters as individual essentialism, there can be no unique set of principles to guide our construction of possible worlds. We can therefore envisage there being standards which conflict over whether or not there genuinely is a world wherein ..., eg. individual a has property φ , according to how liberal they are over what it is to be a . While admitting that $\Diamond\varphi a$ is not determinately either true or false, the view under consideration would still hold to $\Diamond\varphi a \vee \neg\Diamond\varphi a$ because, by whichever standard you chose, there will be a world in which a has φ or there won't.

It should be apparent that this form of argument will justify classical logic even if what is at stake is a highly subjective or evaluative matter. Thus: "Either this (work of art, course of action, ...) is good or it isn't" is assertible, because however you qualify or relativise the predicate one or other disjunct will come out true. We would often find it highly objectionable to be confronted with insistence on LEM in a case like this - with, for example, a morally complex situation which has both good and bad aspects. In forcing us to choose between goodness and nongoodness, it forces us either to misdescribe the situation by coming down in favour of one side at the

expense of the other, or to assert "Well, it is and it isn't", which conveys no information - other than that we in effect decline to describe the situation in the terms offered. (Against the use of classical logic with a borderline vague case, again it may be protested that it does an injustice to our way of describing things. Here it is for the different reason that this idea of sharpening the application of a term, even if temporarily made for the duration of a particular context, does not correspond to our actual linguistic practice. Rather we acknowledge that it can neither be described as red nor as not red, by hedging the description, eg. "It is redish", or even, "It is (an) orangey-red".) Finally, for someone upholding the equation of truth with provability in modal discourse and who takes seriously the fact that our criteria for deciding modal issues may not enable certain questions to be conclusively settled, the use of LEM would seem equally objectionable. The general ethos of "Well, whichever sense you pick, its going to come out one way or the other" is objectionable here, as in the other cases, because of the idea of picking one sense (or relativisation or whatever) at the expense of others². Surely the anti-realist would not find this form of argument acceptable. For someone might argue in the following manner about Jones. Either there is some evidence around concerning his ability with languages or there isn't - the anti-realist did not challenge this instance of LEM. If the former, then there is something which we would count - if we found it - as settling the truth-value of "Jones was good at learning languages". If the latter, it would just count as a second kind of falsity for that sentence. So whichever way the

evidence turns out, the sentence will either be true or false. The whole point, for the anti-realist, is that the difference between the evidence and the no evidence case will not be one we are capable of recognising, so the appeal to the divide between the two cases is illicit in that it could not relate to the meaning we could have conferred on the predicate³. The same goes for someone who would make a more direct and simplistic appeal to the supervenational idea: however we were to complete the meaning of the phrase...

Let us leave this theoretical amalgam, still wanting coherent articulation, and turn to the other combination which Dummett considers. In its modal version, the ontological picture is akin to Lewis's, a belief that possibility forms a further, fifth dimension through which are arrayed the countless possible worlds. Their independence from human cognition is taken to be significant: because nothing travels through this fifth dimension, unactualised worlds play no role in our procedures for determining the truth values of modal statements. Since the ontology of worlds is combined with anti-realist views about meaning which relate it to these procedures, ie. the conditions we are taught to accept and reject assertions under, bivalence is not endorsed for the modal part of our language. Since we are assuming free inter-translation between statements involving modal operators and the corresponding ones with world quantifiers it must be that there are classical truth conditions for modal statements, it's just that the meanings we have conferred upon these statements do not consist in or relate to a grasp of them.

At first sight, it might seem that anyone wanting to

maintain this combination would be confused. He wants to play God yet simultaneously assert that the rest of us cannot attain His perspective. How could he come to believe this? This is not just a psychological query: the obvious route to a thesis about the nature of possible worlds is via one about more normal modal discourse and a construal of the latter in terms of worlds. Bivalence for normal modal discourse would then be a prerequisite for regarding worlds as determinate. But suppose this problem is surmountable - the theorist perhaps takes the Everett/Wheeler 'many worlds' interpretation of QM as good evidence of their independent existence. The problem of coherence remains: what stops these worlds rendering our modal judgements true or false, if the latter are about the former? To turn it around: it looks as if the worlds of QM are not the worlds of modal semantics. The anti-realist argument did not challenge the notion of truth in virtue of something in the world, it will be complained, but rather the general insistence on that something's being there, for every statement, rendering either it or its negation true.

It will help to bring out the incoherence of the position if we return to the idea of God's language, the advantage of which being, as before, its making possible the appeal to principles of translation. For instance, reformulating the position Dummett gives in connection with mathematics, we have a thesis about two languages. Since the theorist believes that natural numbers are "objective, independently existing abstract objects, to each of which the predicate $P(x)$ determinately either applies or does not apply" ([32], p.383), he must believe that in God's language $\exists xP(x)$ is determinately true or false. At the same time he

denies this for our language, convinced by the anti-realist attack on recognition-transcendent truth conditions. This involves a direct assault on the notion of truth which had been agreed upon, in particular its component 1). The problem with this is how we could give direct expression to a counterexample to the conditional 1): if the anti-realist account of our language is right, we cannot presume to understand what God expresses by $\exists xP(x)$ in the way that he does. If God could speak, we would not understand him (aright).

Nevertheless there is something right about this position for the realist who, as mentioned earlier, wishes to emphasise that a reality external to us might transcend our capacities to determine it. Even if we cannot coherently formulate a direct counterexample to 1), the (epistemic) possibility of one arising cannot be ruled out. Recall that at the beginning of this section realism was characterised in terms of three theses. This leaves room for a position calling itself realist in refusing to make truth an epistemic notion, ie. cleaving to its possible verification transcendence, but dropping the metaphysical insistence that there must be something in virtue of which any statement or its negation is true, and thereby also dropping bivalence and LEM. It is a position - renouncing actual verification transcendence (in the presence of non-effectively decidable sentences) - which is reached by McDowell in [67] (see especially section 6. For someone wanting to maintain a truth-conditional theory of meaning, we have seen that there are independent reasons - from consideration of eg. vague predicates - for shedding no tears at the parting of bivalence and LEM.)

A move appealing to a realist, illustrative of his position, is to imagine beings with lesser epistemic abilities than our own, and then to point out that just as it would be foolhardy to base conclusions about the nature of reality on their impoverished conceptions, so it would be in cases where the anti-realist aims his attack at us (the 'flies to wanton boys' move). Consider as an example Poincaré's illustration of the conventionalist theory of geometry. He imagined a Euclidean two-dimensional world which consisted in a circular disk, the temperature of which varied uniformly from absolute zero at the perimeter to a degree of heat at the centre. Its inhabitants, wishing to determine the geometry of their space, start out with rigid rods that contract in direct proportion to reduction in temperature. The data collected by means of this technique would naturally lead them to infer that their world is Lobachevskian and infinite in size - unless, of course, they hit on the idea of the uniform distortion in their measuring apparatus. This challenges the implicit assumption that the rods remain invariant in length when transported. Given this idea, finite Euclidean space would be an equally natural explanation. This exhibits the conventionalist thesis, in that according to Poincaré, from their standpoint only convention could decide which of the rival geometries should be adopted. But notice that, as the example is set up, we have perfect information: we know that their world is finite and has a distorting field. I don't know whether Poincaré intended to include a fact of the matter. Perhaps the conventionalist should abstain from commenting on that issue - the important point for him being that for them, as for us in the four dimensional

case, the choice is a matter of convention.

The fact that there is no possible datum that could decide between two theories does not entail, without further argument, the fact that there could be no such datum. If the speculation about distorting fields could be given empirical content, the choice between the two theories in Poincaré's world would become non-conventional. If the inhabitants of the world could control and measure temperature, ie. if they are not just passively subjected to it, they could measure the elasticity of their measuring rods and the temperature gradient of their world. Until such a point, nothing they had put into the meaning of the theories determined any observable difference in the two, but at that point they should be able to ascertain that they inhabit the kind of possible world described by the second theory. Of course, it would always be open for someone to speculate about another kind of distorting field, but again, until that is given content, the theories would remain on a par with respect to the possible evidence. This is a standard empiricist reply to Poincaré: that there is really just one theory with two different formulations. If a theory is just a device for correlating observations, this is obviously correct, but such holism would be unacceptable to an anti-realist such as Dummett, for whom the basic unit of meaning is the sentence. There is a difference between the theories, but until the hypothesis about distortion is given empirical content, we cannot ascribe to them a proper conception of what that difference is, and therefore, of what their world is really like. Nevertheless, we can see that there is a difference there, a difference which at one time transcended their

powers of ascertainment. Likewise in our case, the view that spacetime is external to us makes it natural to believe that we do inhabit a definite possible world even if we cannot fully conceive how it is. We have no guarantee that it does transcend our ability to ascertain its nature, but nor can we be sure that it does not.

This leaves us with a position in which metaphysical or ontological questions cannot be decisively settled by consultation of the theory of meaning for our language. For example, we have no criterion by means of which to decide whether or not some aspect of reality - possible worlds, for instance - is external to us. This seems the correct position to me - theories about the nature of reality which are metaphysical in the sense of transcending the empirical are not capable of being settled conclusively. As Prior says, "In doing metaphysics there is still no substitute for 'the choice of the soul'; or, if you like, prejudice" ([84], p.93).

CONCLUSION

Someone who was looking for modality to provide something of a 'crucial experiment' to decide between rival semantic programmes might find the preceeding considerations a little disappointing, in that none of the three programmes discussed decisively fails to accomodate modal locutions. However we need not look for outright failure as the point of comparison between the theories so much as the manner in which they treat modality. In this conclusion I want to draw together the strands as far as possible by comparing the theories with regard to the kind of light they would have us regard modality in.

I would start by mentioning a measure of agreement between the Davidsonian and Montaguean approaches, in their distinguishing between objective modalities of a broadly logical or metaphysical nature and their epistemic counterparts. For the Davidsonian, the difference can be illustrated in the following way, based on the general case of interpreting one language within another:

- (1) Wahr(Maybe there is a married bachelor, englisch) =
 Vielleicht Wahr(there is a married bachelor, englisch).

As far as the average German speaker is concerned, it may be that the sentence quoted on the RHS is true. We know better of course; the sentence quoted on the LHS, and hence the whole of the LHS, is false. This suggests that epistemic modalities, because prey to the Wallace argument, cannot be given the homophonic treatment of their objective alethic counterparts. On the Montaguean approach, epistemic modality is recalcitrant in not permitting substitution on the basis of equivalence in all possible worlds (§2.4).

However, the agreement on this point masks a fairly crucial disagreement between the two schools.

On the natural account of the way the model theorist works (§2.1), a language is something whose expressions by and large could have had other meanings. But according to the standard justification for the use of broadly logical \Box to interpret itself, it is of the essence of a language that its expressions have the meanings they do. A conception of language cannot simply be adopted or dropped as a matter of expedience, according to the kind of semantic theorising we are engaged in at the time. From the truth theorist's point of view, if $\Box\varphi$ is true in L , $T(\varphi, L)$ must be regarded as strictly necessary. That could not be if we were also free to adopt the alternative conception of L .

Should we perhaps resolve the difference in favour of the truth theorist's conception of language? After all, on the PTQ style of semantics, the only language which receives a direct model-theoretic interpretation, and which therefore requires the conception which accompanies it, is the artificial language of tensed intensional logic, not English itself. The model-theoretic conception is suited to artificial languages, since typically the only expressions of such a language to have a definite meaning bestowed on them are anyway the ones whose interpretation is held constant from model to model.

But we are by no means in a position to dismiss the alternative conception of language. What about the kind of fact which Kripke cites, such as that Hesperus might have picked out a different planet to the one it in fact does? In fact what Kripke says about this example goes against

the conception - namely that such a usage would involve a different language ([60], p.91). Indeed that is a common enough thought - that to change the meaning of even a single word is 'to speak another language'. But it is not compulsory - we could reformulate it as 'to use the language in a different way'. Indeed it seems odd to make the identity of a whole language depend on the meaning of a single word. Part-whole essentialism becomes less plausible the more parts there are - it makes too much turn on too little. There is a spectrum ranging from artifacts fashioned from a single lump of matter to those of the order of the ship of Theseus. Its unsuitability in the case of natural language can be illustrated by the fact mentioned by Dummett ([29], pp.584-85), that speakers are unlikely to agree on the definition of every word, ie. on the sense they attach to them. (Most dramatically: no two speakers are likely to share exactly the same lexicon of proper names.) In consequence, there would be no shared language but a multitude, one to each speaker.

Is the alternate conceivability of the language under study enough to invalidate the truth theorist's homophonic treatment of modality? Here we have to distinguish between saying of the language which is the OL of the semantical theory that we can choose to conceive it in different ways, and saying that the choice is between different (kinds of) languages to serve as the OL of the theory. The former option is not consistent with an S4-like conception of necessity (since there would be some facts about the language, English, which were necessary but not necessarily necessary), yet S4, regardless of its other merits, is virtually indispensable to the theory (cp.[48],

p.451). Rather the truth theorist takes the latter option, of choosing a language which could not be conceived otherwise (what is conceivable is another language, similar to it except that it bears its semantical attributes contingently). Both the model theorist and the truth theorist study object languages which are proxies for English; the difference between them discussed here lies in the different natures of these proxies.

Of course, this also leaves the truth theorist with the options of treating alethic modality using either possible worlds, like the model theorist, or the methods of parataxis.¹ To some extent these treatments raise the question of the place of analysis in semantic theory. Possible worlds, for instance, because, as I tried to show in §1.5, talk of them represents an increase in expressive power over plain modal operator talk, albeit a gradual one. It is generally agreed that model theory is better suited to the business of analysis than truth theory (eg. by systematising its results, as in §2.3), and of course in aiming for homophony the truth theorist wants to leave the conceptual resources of the OL untouched as far as possible (§1.1). However Davidson does not rule it out; he proposes that analysis, at least of particular expressions, should be "preceded by or at any rate accompanied by" an account of logical form ([13], p.317). But for kinds of expressions, or particular constructions, such as quotation, it is unlikely that an account of form could ever properly precede some sort of analysis. Consider by way of example Davidson's spelling account of quotation. It is a technically adequate account of form, in that it permits us to derive the reference of an infinity of quotations. But

whether it could serve as a basis for an analysis of what is going on in quotation is another matter. Illiterate people can quote the words of another, by attending to the words they hear; clearly then, at least as far as mastery is concerned, spelling is not necessary for quotation. It is an overdiscernment of structure - in general, we only want to cut as fine as the words, not the letters which comprise them. (As such, of course, someone could reply that it would be ruled against by Davidson's own criterion C) (from §1.1).)

Aside from this (bland point about analysis informing our selection of form proposals), there remains a doubt as to whether certain accounts of logical form are not artifacts arising from the need to mould sentences into a form suitable for processing by truth theory. Consider the paratactic theory. In its original, non-quotational version it does not seem very attractive - the claim that that functions as a perfectly ordinary demonstrative, for instance, is belied by the fact that propositional-level phrases like the belief that have no individual concept level correlate (I use the terminology of §2.2), eg. the man that. (Where that, of course, does not introduce a subordinate clause but should be acting as a demonstrative.) Indeed we cannot use belief that etc. in a demonstrative manner either, as the Wilde analogy would suggest, eg. there is a possibility that (cp. 2.2.15), or the thought that amuses Bill (instead we would use the qualified demonstrative: that thought amuses Bill). So let us turn to the quotational version distinguished in §1.3 and ask of it: does it provide the basis for an analysis of intensional contexts?

It is important to distinguish several different aspects of quotational treatments of these contexts. Firstly, from a philosophical point of view a quotational treatment is always possible. As Geach has put it

oratio recta can be used metaphorically to report what somebody thought, 'said in his heart' ... such constructions are frequent in the Authorised Version of the Bible: "The fool hath said in his heart 'There is no God'" ... Clearly we could always describe judgments in this way; oratio obliqua is logically superfluous. ([43], p.80)

Secondly, a quotational treatment is apparently required by the presence of certain phenomena. In [116], Barbara Partee classified propositional attitude verbs according to what they permitted by way of substitution within their scope. At one end of the scale were placed the 'emotive' verbs like regret that, be surprised that, for which logical equivalence suffices. At the other were the verbs of communication such whisper that for which, since they "emphasise the manner of communication, it is not surprising that the form as well as the content of the embedded clause is significant" ([116], p.326). So even if we had a stable notion of proposition, certain of these verbs would not permit substitution of content sentences on the basis of their expressing the same proposition, being quasi-quotational in nature (recall the 'green ink' example of §1.2). Moreover, there are certain phenomena from within the content sentence, as it were, which call out for some kind of quotational treatment (cp. §2.4). They therefore call into question somewhat Partee's scale of propositional attitude verbs. Of Kripke's peripatetic Pierre we might wish to say that he regretted that he had come to London, but not that he regretted that he had come to Londres. On one natural way of describing the situation,

he couldn't regret that, since he wasn't aware of it. But this does not affect Partee's central conclusion, that we should see propositional attitude verbs as sentence-taking rather than proposition-taking ([116], p.335). Those verbs one might have taken to be proposition-taking would really be sentence-taking verbs which (by and large) allowed substitution on the basis of expression of the same proposition. The point is already suggested by Geach's "said in his heart", ie. "said something amounting to this:... ", with varying standards of 'amounting to'. The point is, though, that once we are dealing with any intensional idiom which is not as objective as strict necessity, there is the danger that substitution on the basis of necessary equivalence will be problematic.

Finally, we can ask whether a quotational treatment will provide an analysis, or explanation, of all the interesting aspects of propositional attitudes. This all depends. It may be, for example, that the explanation of the failure of substitutivity salva veritate of necessarily equivalent expressions of a certain type within these contexts involves a quotational element. For even if two names are as close to synonymous as one could hope to find, they will not be everywhere intersubstitutable, as the following construction of Mates's serves to remind us. If φ and φ' differ only in that each contains one of the pair, then of the following two sentences, the first will be true and the second false:

- (2) Nobody doubts that whoever believes that φ believes that φ .
- (3) Nobody doubts that whoever believes that φ believes that φ' .

In other words, if believe was proposition-taking, we would have two synonymous sentences with differing truth values. One reason someone could have for entertaining the doubt expressed in 3) is that they are thinking of failures of competence (eg. in a person who only properly understood one of the names). But this need not be the case, as the examples of §2.4 show. It appears that their simply being different expressions is sufficient for their non-substitutivity salva veritate in these contexts. With other kinds of expressions, eg. equivalent logical or mathematical expressions, this is not the explanation we would look for. Here it is not that they are different expressions per se which is important but rather the fact that they are constructed in different ways out of different units. This is obviously not a point which needs the adoption of a quotational point of view for its appreciation.

A quotational treatment certainly should not be taken to be an analysis of propositional attitude contexts in the sense of proposing that we should take the particular sentence quoted as something the subject of the attitude is directly related to. As has often been pointed out, he need have no inkling of it (cp. eg. p.42 above). To the extent that we want to say that animals have beliefs, for example - and surely we do want to say that the dog expects that it's about to be taken out - we cannot treat these as attitudes directed towards sentences.

Having found this area of convergence between the Davidsonian and Montaguean approaches - something apparently forced on them by the nature of the subject matter, psychological modality itself - let us turn to compare them

with their anti-realist rival. One kind of realist is a semantic theorist who, if only implicitly, adheres to the principle of bivalence. In undecidable cases the only means to support this adherence would appear to lie in the appropriate conception of a determinate reality. In effect, the theorist becomes committed to a certain metaphysical conception, according to the anti-realist. It is commitment without support, since there is no further route to support the metaphysical picture. Alternatively, we may consider a realist who is primarily a metaphysician and who, if he engaged in semantic ascent, would find himself committed to bivalence for the realm of discourse concerned. Either way it is with the principle of bivalence that the anti-realist's critique begins, with its failure to square with the facts of language usage. But there are different informal formulations of that principle. It may well be that the latter kind of realist cannot quibble with the occurrence of 'determinately' in the form of the principle he is committed to, as in 'Every sentence is either determinately true or determinately false', but there need be nothing in the kind of truth definition discussed in the first two chapters which commits them to this rather than the simpler formulation which omits the word. There is an even weaker principle, 'Every sentence can be regarded as either true or false', which might be thought of as a form of bivalence, at work in certain theories, in that these theories, while validating classical reasoning, do not rely on the previous, standard formulation of bivalence. A borderline vague sentence, according to such a theory, can be regarded as true or false because, for all practical purposes, we can set the standards within any particular

context which will decide which way it is to be counted.

How does this affect the semantics of modal notions? Let us start with dispositions. In §3.2 we saw that there is a case for allowing that sentences involving certain disposition terms can be regarded as either true or false even though it may be beyond the current means of the language community employing them to determine which. Take the idea of a soluble thing as that which would dissolve if it were immersed. What is it for the latter to hold; what sustains that subjunctive? Recalling Carnap's rather elementary model, we need to know how to progress, in our application of the predicate, from the immersed to the non-immersed. It was argued in §3.2 that what allows the application of the form of bivalence lately distinguished was, to put it simply, the possibility of there naturally existing a distinction which it is for the community to try to uncover. This kind of realism is of course familiar from the work of Kripke, Putnam and others on natural kind terms. In normal circumstances our ordinary, phenomenal criteria for distinguishing tigers, say, do not come into conflict with a classification of them according to a genuinely scientific account of what those creatures are. But our ordinary, phenomenal stereotype does not constitute a definition of 'tiger'. More importantly, we hold our usage accountable to the findings of the scientific account. Tigers are 'that kind of creature' (cp.[59], p. 319), where this is not defined in advance as whatever is quadrupedal, yellow with black stripes, etc..., (compare the example of fool's gold); it is for the scientist to discover if he can what constitutes that kind of animal.

One thing this points to is the fact that the

principle 'meaning is use' is not as uncontroversial as Dummett would have us think. I merely note here that a case can be made against it (for Putnam himself has recently ([86], Part 3) rejected the idea that his earlier work poses a challenge to it). The principle is naturally allied to an emphasis on the person-to-person, communicative side of language, and fits naturally with a certain analogy with chess pieces and the moves they make (cp. [33] p. 216). But words are not solely like chess pieces, which can be moved around without accountability to an external reality. Words are not mere pawns; it is the function of some to make reference to the real world. A simple illustration of the challenge posed is provided by Kripke's examples of our use of a name being accountable to the referent at the end of the causal chain, whatever that might turn out to be, rather than the 'meaning' extractable from current usage, ie. the property commonly associated with the name (cp. the discussion of Gödel, [59], p.294).

If this is along the right lines, then such expressions, including terms for physical dispositions, do provide something of a test case for the realist and anti-realist, allowing the former to meet the latter's challenge in a limited but fairly direct manner. The realist could say that the meaning of the term is not necessarily determined by its current usage, ie. may transcend in a restricted way the conditions of application so far worked out by the community, or he could say that its contribution to truth conditions may transcend the (usage-determined) meaning it has. However, it was noted at the end of §3.2 that psychological dispositions do not furnish such a clear test. Here usage does not exhibit to the same extent the

assumption that there is some underlying property as yet only partially uncovered - obviously a precondition for saying that usage is held accountable to such a property.

Where there is an external reality, then, there is hope for realist assumptions to be borne out. Where there is none, *ex hypothesi* there can be no natural distinction for our terms to aim to relate to. This is how it appears to be with our use of ordinary modal words - for example, in relation to the postulated domain of possible worlds. There can be no empirical proof of their existence, since any 'world' we have access to must, by definition, be part of actuality. The best-known attempted non-empirical proof, Lewis's much discussed section 4.1 of [63], is not as convincing as might first appear. For instance, in his attempt to show the ordinary speaker's commitment to the existence of possible worlds, Lewis construes 'there are many ways things could have been' as a quantification over 'ways things could have been'. A better reading of the expression, however, is as a quantification over 'ways', ie. properties, of 'things', ie. such that things could have been those ways rather than the way they are. (Stalnaker has made a similar point.) Commitment to properties is less controversially a feature of our ordinary talk. While possible worlds can provide a useful *façon de parler*, their reification does not illuminate the nature of modality. Whereas in a discipline like arithmetic, proof of necessity is more or less conclusive, this is not the case with the conceptual truths of philosophy. In such a discipline there is no conclusive verification of what must hold in every possible world. If we have no firm reason for overturning the commonsense belief that there is

no such external modal reality, the line of defence considered for the realist will not be available.

SOME NOTATION

Even if a reconciliation in the content of the theories were possible, a reconciliation of their notational styles is not. I have tried to cobble together a notation and apply it uniformly, though this has led to some characters playing two or more roles. Here are some of the more notable ones:

- T - the most important semantical predicate, of truth, which is subjected to a number of different relativisations, signalled by different subscripts. T_m , for example, is the predicate of truth relative to a model. T is also used to denote the set of moments of time, t.
- R - complementing T, and relativised in the same style, "the reference of ..."
- TR - expresses the translation relation.
- A - the domain of entities (possibly: possible entities). A also figures as an actuality operator.
- W - the set of possible worlds, w.
- @ - a constant denoting the actual world.
- - expresses necessity (itself the theme for several variations).
- ◇ - expresses possibility.
- P - "it was the case that ..."
- F - "it will be that ..." But F is also a variable over language fragments (Ch. 1), an assignment function (§2.1) and stands for various syntactic operations (§2.2).
- - the subjunctive conditional (not necessarily contrary-to-fact).

Quotation:

- ~~~~ - is a means of quasi-quotation; so eg. necessarily φ = the result of writing "necessarily" followed by φ .
The typist's answer to the boldface used in Montague Grammar, it is typically the quotation a theorist would make of the expressions of the OL he is studying.
- ' - single quotes are either quotation marks belonging to the OL, or scare quotes.
- " - double quotes are a general, all-purpose device of quotation. When quoting other authors, I have sometimes taken liberties in adapting their notation to

mine.

To further confusion, autonymy is also used (not to mention mentioned).

Abbreviations:

"OL" for "object language", "ML" for "metalanguage",
 "LHS" for "left hand side", "RHS" for "right hand side",
 "L.E.M." for "the law of excluded middle".

Numbering:

a number thus: "12)" refers to the example numbered
 twelfth in the section in which it occurs; thus: "1.3.7"
 to the seventh example of section 1.3; thus: "[2]" to
 the second entry in the bibliography.

SECTION 1.3

[25], p.17, is both propositional attitudes in
 general and also the various kinds of sentences.

16. Following Davidson's paradigmatic treatment of mood -
 see [25], p.1.

"pragmatic" in a sense which is not identical to the
 application of truth theory - see [17], 28.4.

[13], p.307, though at [15], pp.147-48 he gives his
 support to the concept of indeterminacy.

see [25], p.18.

An example of the latter half of this dilemma would be
 the point made by Millard, in [7], p.203, that
 Davidson's analysis of propositional attitudes will
 license the problematic substitution.

It is therefore assumed that a truth theory will be
 suggested by a translation manual satisfying the
 pairings required by this rule.

SECTION 1.4

In [20], Wallace points that 1.2.3 might work for some
 special cases, but doesn't say which. In [20], he
 promotes the use, in the ML, of an operator "it is a
 matter of meaning since that".

[109], p.255, or even to use the comparison
 with attitudes, [110], p.107.

Holden also has difficulties understanding this term.

FOOTNOTESSECTION 1.1

- 1 [19], p.84, whereat a neat summary of this first formu-
lation of the Davidsonian project is to be found.
- 2 I ignore certain complications concerning tense.
- 3 This seems to be the setup envisaged by Wiggins in
sections V and VI of his [111].
- 4 Compare here Quine's remarks in [91].
- 5 [30], p.105. It will be clear that I disagree with
Dummett's stated reason for wondering this.

SECTION 1.3

- 1 [25], p.17, ie. both propositional attitude contexts in
general and also the various moods of sentences.
- 2 ie. following Davidson's paratactic treatment of mood -
see [25], p.18.
- 3 "pragmatic" in a sense which is not inimical to the
application of truth theory - cp. [17], fn.4.
- 4 [13], p.307, though at [15], pp.147-48 he gives his
support to the concept of indeterminacy.
- 5 cp. [25], p.18.
- 6 An example of the latter half of this dilemma would be
the point made by McGinn, in [71], p.203, that
Baldwin's analysis of propositional attitudes will
license the problematic substitutions.
- 7 It is therefore assumed that a truth theory will be
augmented by a translation manual delivering the
pairings required by this rule.

SECTION 1.4

- 1 In [107], Wallace hints that 1.2.3 might work for some
special cases, but doesn't say which. In [108], he
promotes the use, in the ML, of an operator "It is a
matter of meaning alone that".
- 2 [109], p.293, my emphasis - see also the comparison
with abilities, [110], p.107.
- 3 Baldwin also has difficulties understanding this term,

[6], p.11. It should be clear that 15) is really only a first stab at conveying property modification.

- 4 It might be suggested that because "satisfy" is a term of art, we have a certain amount of freedom in what we understand by "noisily satisfy" - so why not say that to noisily satisfy eat an apple just is to noisily eat an apple? But this ignores the fact that properly speaking the predicate to be modified in the ML is "be true in L relative to s", for which this reading is less plausible.

- 5 Cp. McDowell's [69], p.221, where he gives the idea a personal twist: two people speak the same language if sufficiently many equiform utterances of theirs mean the same.

SECTION 1.5

- 1 [19], p.83. All this under the proviso that we wish to account for certain inferences as valid.

SECTION 2.1

- 1 Cp. the kind of criticism voiced by Strawson, in [97], and cp. [16], p. 759.
- 2 The idea of models as possible dictionaries for a language is especially clear in IMT; cp. [73], pp.209-10.

SECTION 2.2

- 1 Or would be, if the PTQ fragment contained adjectives; cp. [73], p.193.
- 2 Plausible because intuitively a possibility is a state of affairs, one satisfying some appended description. A state of affairs can be thought of as a set of worlds, as can our propositions p.
- 3 Plausible because the it is semantically inert - it doesn't refer to anything, and in that ϕ δ we have something already syntactically and semantically complete. The practical value of placing the 'modality' δ at the front is illustrated by the very sentence in the text - with a very long that-clause

the object of the verb would be hard to ascertain if the order was reversed.

SECTION 2.3

- 1 This thought seems to be lying behind Wiggins' words:
 "The point where our intuitions begin to falter about the notion of necessity which we express by [necessarily] is, I think, just beyond the frontier marked by Lemmon's system S0.5" ([109], p.312).
- 2 Iteration in natural language is less of a clue to our ordinary conception of necessity than the usual systems of modal logic might lead us to believe. In ordinary speech the repetition of a word may simply be pleonastic. For unambiguous amplification we use words like very, absolutely, etc.
- 3 "I shall investigate the sort of constraints that would have to be placed on the alternativeness relation ... [reflecting] the natural or naive modal concepts of ordinary language discourse" ([57], p.3).
- 4 The must form "has an implication that [the generality] results from an essential characteristic" ([75], p.155).
- 5 Cp. [73], p.112, an idea originally proposed by Scott, investigated since by Aqvist, Burgess, Chellas and Kamp.
- 6 If instead of 26) we had, say, "That is not an option for me", we would be less inclined to think that there was a shift of sense involved, but the dispute would still have arisen.

SECTION 2.4

- 1 What Montague actually accepted was a weakening of the principle which Carnap had originally adopted, of taking assent to be only non-conclusive evidence of belief.
- 2 Dummett attributes it to Kripke, [29], pp.112-13. Cp. our discussion of Kripke on yellow, §3.2 below.
- 3 Whether all description theorists would accept this dichotomy is another matter. Those who see the relation between description and name as criterial

presumably would not. For the notion of criterion, see eg. [4], though for Kripke's understanding of it cp. [62], fn.3.

SECTION 3.1

- 1 Dummett acknowledges a connection between 1) and his principle C) (cp. §3.4 below): "if it were in principle impossible to know the truth of some true statement, how could there be anything which made that statement true?" ([31], p.99).
- 2 And compare the comparisons of McDowell, [67], section 11, and McGinn, [70], pp.32-33.
- 3 "in point of meaning ... a word may be said to be determined to whatever extent the truth or falsehood of its contexts is determined" ([87], p.89, my emphasis).

SECTION 3.2

- 1 Cp. [33], p.224, and also the example Lewis gives, [63] p.80, of 8) being consistent with the denial of 2), where there is no correlate of χ , of 4) and 5).
- 2 [72], p.55 - a contrast Mellor goes on to repudiate.
- 3 In this case we have what Carnap calls a bilateral reduction sentence, ie. $\varphi \rightarrow (\psi \equiv \chi)$.
- 4 Cp. Mellor's example of nuclear safety precautions, [72] p.62.
- 5 Cp. Sartre's discussion of the paederast, [94], Pt.1, Ch. 2, section III.

SECTION 3.3

- 1 Of course, if sense is equated with extension in every possible world, imparting any (non-trivial) necessary truth will count as imparting sense. Although Kripke may seem committed to such an equation (cp. §2.4), what he says about " π " on p.278 of [59] shows that he rejects it (although given his stipulation of analytic = necessary + a priori (fn.21), it's not clear why.)
- 2 "A possible world is given by the descriptive conditions

- we associate with it" (Kripke, [59], p.267).
- 3 That maximal consistency embodies a classical, realist conception of worlds has been pointed out by, amongst others, G. Baker in [4].
- 4 If philosophy is "the science of the possible" (Russell), philosophers should appreciate this from the tentativeness of their own conclusions. The kind of conceptual issue philosophers are concerned with - individual essentialism, for instance - are by their nature not susceptible to conclusive verification. But even if we suppose such a doctrine true, it is still worth pointing out that it would not necessarily render determinate every matter of possibility. Suppose, for example, it is claimed essential to Nixon that he came from the sperm and egg he in fact came from. A question we may ask is: how much freedom is there to tamper - in thought - with these gametes? Eye colour is a genetically determined trait. If we are prepared to grant truth to the supposition that Nixon's eyes could have been a different colour, not only do we countenance Nixon's potentiality to have been different but, ceteris paribus, we also recognise the very gametes from whence he sprang to have that right. But surely the identity of a gamete will not survive a hypothetical total enucleation and replacement of genetic material. So how much change is permissible? Cp. Wiggins on general essentialism, [110], p.124.
- 5 "the expression 'a direct insight into counterfactual reality' provides no picture of what these powers [of] observation consist in" (Dummett, [31], p.100).

SECTION 3.4

- 1 "Against the fundamental intuition of Verificationism... the contrary intuition of realism: that human experience is only a part of reality, that reality is not part or whole of human experience" (Putnam, [85], p. 273).
- 2 Which would make a possible exception of the past - on the assumption that it is determinate that there was just one actual past history, even if it is

indeterminate which.

- 3 As Dummett says, there is no challenge to the principle that "there can be no circumstances in which a statement can be recognised as being, irrevocably, neither true nor false" ([33], p.xxx).

CONCLUSION

- 1 Perhaps this is a good point to clarify my attitude to Gupta. It is implicit in my discussion of his work (p.56 above) that Gupta can explain his truth predicate T_2 for the homophonic treatment of \Box by appeal to ordinary English, eg. as 'truth relative to what the sentence actually means'. But this conflicts to some extent with the case in §1.5 for the redundancy of the ordinary actually - the significance here being that it is not strong enough to secure the truth of $\Box T_2(2+2=4, \text{English})$, say (since "'2+2=4' might actually have meant something false in English"). So on this construal we must stick to the more explicit 'actual world' idiom which Gupta himself adopts.

BIBLIOGRAPHY

Page references are to the reprinted version of an article, when mentioned. With Dummett's and Montague's well-known collections of essays [33] and [73], I have not differentiated articles occurring therein, with the exception of using the common abbreviation "PTQ" for Montague's "The Proper Treatment of Quantification in Ordinary English", chapter 8 of [73]. Two abbreviations used below: JPL - The Journal of Philosophical Logic, PASSV - Proceedings of the Aristotelean Society, Supplementary Volume.

1. Adams, R.M.: "Theories of Actuality", Nous 8 (1974), pp.211-31.
2. Austin, J.L.: "Truth", reprinted in Philosophical Papers. Oxford: Oxford University Press, (1961).
3. ---- "Ifs and Cans", reprinted *ibid*.
4. Baker, G.: "Criteria: A New Foundation For Semantics", Ratio XVI (1974), pp.156-89.
5. Baldwin, Thomas: "Quantification, modality, and indirect speech", in Blackburn, S. (ed.), Meaning, Reference and Necessity. Cambridge: Cambridge University Press, (1975).
6. ---- "Wiggins and the De Re 'Must'", in Dancy, J. (ed.), Papers on Language and Logic. Published by Keele University, (1980).
7. Bennett, Michael: "A Variation and Extension of a Montague Fragment of English", in [78].
8. Burge, Tyler: "Self-Reference and Translation", in [47].
9. Burgess, John P.: "The Unreal Future", Theoria XLIV (1978), pp.157-79.
10. Carnap, R.: "Testability and Meaning", Philosophy of Science 3 (1936), reprinted in part in [103].
11. Chellas, B.F.: Modal Logic. Cambridge: Cambridge University Press, (1980).
12. D'Alessio, J.C.: "Dispositions, Reduction Sentences and Causal Conditionals", Critica 1 (1967), pp.65-76.
13. Davidson, Donald: "Truth and Meaning", Synthese 17 (1967), pp.304-23.

14. ---- "The Logical Form of Action Sentences", reprinted in [27].
15. ---- "On Saying That", reprinted in [27].
16. ---- "True to the Facts", The Journal of Philosophy LXVI (1969), pp.748-64.
17. ---- "Semantics for Natural Languages", reprinted in [27].
18. ---- "Action and Reaction", Inquiry 13 (1970), pp. 140-48.
19. ---- "In Defense of Convention T", in Leblanc, H. (ed.), Truth, Syntax and Modality. Amsterdam: North-Holland, (1973).
20. ---- "Radical Interpretation", Dialectica 27 (1973), pp.313-28.
21. ---- "Belief and the Basis of Meaning", Synthese 27 (1974), pp.309-23.
22. ---- "Reply to Foster", in [40].
23. ---- "The Method of Truth in Metaphysics", in French et al. (eds.), Midwest Studies in Philosophy II (1977).
24. ---- "Reality Without Reference", Dialectica 31 (1977), reprinted in [82].
25. ---- "Moods and Performances", in [66].
26. Davidson, D. and Harman, G. (eds.): Semantics of Natural Language. Dordrecht: Reidel, (1972).
27. ---- The Logic of Grammar. Encino: Dickenson, (1975).
28. Delacruz, E.B.: "Factives and Propositional Level Constructions in Montague Grammar", in [78].
29. Dummett, M.A.E.: Frege: Philosophy of Language. London: Duckworth, (1973).
30. ---- "What is a Theory of Meaning?", in Guttenplan, S. (ed.), Mind and Language. Oxford: Clarendon, (1975).
31. ---- "What is a Theory of Meaning?(II)", in [40].
32. ---- Elements of Intuitionism. Oxford: Clarendon, (1977).
33. ---- Truth and Other Enigmas. London: Duckworth, (1978).
34. ---- "What Does The Appeal To Use Do For The Theory Of Meaning?", in [66].
35. ---- "Comments", in [66].
36. ---- "Comments on Professor Prawitz's Paper", in von

- Wright, G.H. (ed.), Logic and Philosophy. The Hague: Nijhoff, (1980).
37. Evans, Gareth: "Semantic Structure and Logical Form", in [40].
 38. ---- "Pronouns, Quantifiers and Relative Clauses", Canadian Journal of Philosophy VII (1977), reprinted in [82].
 39. ---- "Reference and Contingency", Monist 62 (1979), pp.161-89.
 40. Evans, G. and McDowell, J. (eds.): Truth and Meaning. Oxford: Oxford University Press, (1976).
 41. Frege, Gottlob: The Foundations of Arithmetic. Oxford: Blackwell, (1950).
 42. ---- "On Sense and Reference", in Geach, P. and Black, M. (eds.), Translations from the Philosophical Writings of Gottlob Frege. Oxford: Blackwell, (1977).
 43. Geach, Peter: Mental Acts. London: Routledge and Keegan Paul, (1957).
 44. Goodman, Nelson: Fact, Fiction and Forecast. Indianapolis: Bobbs-Merrill, (1973).
 45. Grandy, Richard E.: "Some Remarks About Logical Form", Nous VII (1974), pp.157-64.
 46. Grice, H. Paul: "Further Notes on Logic and Conversation", in Cole, P. (ed.), Syntax and Semantics, Vol. 9. New York: Academic Press, (1978).
 47. Guenther, F. and Guenther-Reutter, M(eds): Meaning and Translation. London: Duckworth, (1978).
 48. Gupta, Anil: "Modal Logic and Truth", JPL 7 (1978), pp.441-72.
 49. Hart, W.D. and McGinn, C.: "Knowledge and Necessity", JPL 5 (1976), pp.205-8.
 50. Hazen, A.: "Expressive Completeness in Modal Language", JPL 5 (1976), pp.25-46.
 51. Hempel, Carl G.: Aspects of Scientific Explanation. London: Collier-Macmillan, (1965).
 52. Honoré, A.M.: "Can and Can't", Mind LXXIII (1964), pp.463-79.
 53. Kamp, J.A.W.: "Two Theories About Adjectives", in Keenan, E.L. (ed.), Formal Semantics of Natural Language. Cambridge: Cambridge University Press, (1975).

54. Kamp, Hans: "The Philosophical Significance of Intensional Logic", PASSV XLIX (1975), pp. 21-44.
55. ---- "Quantification and Reference in Modal and Tense Logic", in Schmidt, S.J. (ed.), Pragmatik/Pragmatics II. Munich: Fink, (1976).
56. Kaplan, D.: "What is Russell's Theory of Descriptions?" reprinted in [27].
57. Karttunen, Lauri: "Possible and Must", in Kimball, J.P. (ed.), Syntax and Semantics, Vol. 1. New York: Seminar Press, (1972).
58. Kratzer, Angelika: "What 'Must' and 'Can' Must and Can Mean", Linguistics and Philosophy 1 (1977), pp. 337-55.
59. Kripke, Saul: "Naming and Necessity", in [26].
60. ---- "Identity and Necessity", reprinted in Schwartz, S.P. (ed.), Naming, Necessity and Natural Kinds. Ithaca: Cornell University Press, (1977).
61. ---- "Is There a Problem About Substitutional Quantification?", in [40].
62. ---- "A Puzzle About Belief", in [66].
63. Lewis, David: Counterfactuals. Oxford: Blackwell, (1973).
64. ---- "Scorekeeping in a Language Game", JPL 8 (1979), pp.339-59.
65. Mackie, J.L.: "Dispositions, Grounds and Causes", Synthese 34 (1977), reprinted in [103].
66. Margalit, A.(ed.): Meaning and Use. Dordrecht: Reidel, (1979).
67. McDowell, John: "Truth Conditions, Bivalence and Verificationism", in [40].
68. ---- "Physicalism and Primitive Denotation: Field on Tarski", Erkenntnis 13 (1978), reprinted in [82].
69. ---- "Quotation and Saying That", in [82].
70. McGinn, Colin: "Truth and Use", in [82].
71. ---- "Operators, Predicates and Truth-theory", in [82].
72. Mellor, D.H.: "In Defense of Dispositions", Philosophical Review 83 (1974), reprinted in [103].
73. Montague, Richard: Formal Philosophy. New Haven: Yale University Press, (1974).
74. Newton-Smith, W.H.: The Structure of Time. London:

- Routledge and Keegan Paul, (1980).
75. Palmer, F.R.: Modality and the English Modals. London: Longman, (1979).
 76. Parsons, K.P.: "Ambiguity and the Truth Definition", Nous VII (1973), pp.379-94.
 77. Parsons, T.: "Modifiers and Quantifiers in Natural Language", (mimeographed).
 78. Partee, Barbara H.: Montague Grammar. New York: Academic Press, (1976).
 79. Peacocke, Christopher: "An Appendix to David Wiggins' 'Note'", in [40].
 80. ---- "Necessity and Truth Theories", JPL 7 (1978), pp.473-500.
 81. ---- "Causal Modalities and Realism", in [82].
 82. Platts, Mark: Reference, Truth and Reality. London: Routledge and Keegan Paul, (1980).
 83. Popper, K.R.: Appendix X to The Logic of Scientific Discovery, London: Hutchinson, (1959), reprinted in part in [103].
 84. Prior, A.N. and Fine, Kit: Worlds, Times and Selves. London: Duckworth, (1977).
 85. Putnam, Hilary: Mind, Language and Reality. (Philosophical Papers, Vol. 2). Cambridge: Cambridge University Press, (1975).
 86. ---- Meaning and the Moral Sciences. London: Routledge and Keegan Paul, (1978).
 87. Quine, W.V.O.: "Truth by Convention", reprinted in The Ways of Paradox, Cambridge: Harvard University Press, (1976).
 88. ---- "Two Dogmas of Empiricism", reprinted in From A Logical Point Of View, New York: Harper and Row, (1961).
 89. ---- "Reference and Modality", reprinted *ibid*.
 90. ---- Word and Object. Cambridge: M.I.T., (1964).
 91. ---- "Comment on Donald Davidson", Synthese 27 (1974), pp.325-29.
 92. ---- "Disposition", an excerpt from Roots of Reference, La Salle: Open Court (1974), reprinted in [103].
 93. Russell, Bertrand: An Inquiry Into Meaning and Truth. Harmondsworth: Penguin, (1973).
 94. Sartre, Jean-Paul: Being and Nothingness. New York:

- Washington Square, (1966).
95. Smith, N. and Wilson, D.: Modern Linguistics. Harmondsworth: Penguin, (1979).
 96. Stalnaker, R.C. and Thomason, R.H.: "Abstraction in First Order Modal Logic", Theoria XXXIV (1968).
 97. Strawson, P.F.: "Truth", PASSV XXIV (1950).
 98. ---- "May Bes and Might Have Beens", in [66].
 99. Tarski, Alfred: "Der Wahrheitsbegriff in den formalisierten Sprachen", Studia Philosophica 1 (1935), pp. 261-405.
 100. ---- "The Concept of Truth in Formalised Languages", in Logic, Semantics and Metamathematics. Oxford: Oxford University Press, (1956).
 101. Taylor, Barry: "Truth-theory for Indexical Languages", in [82].
 102. Thomason, Richmond: "Necessity, Quotation and Truth: An Indexical Theory", in Kasher, Asa (ed.), Language in Focus. Dordrecht: Reidel, (1976).
 103. Tuomela, R. (ed.): Dispositions. Dordrecht: Reidel, (1978).
 104. Van Benthem, J.F.A.K.: "Tense Logic and Standard Logic" Logique et Analyse 20 (1977), pp.395-437.
 105. Vendler, Zeno: Linguistics in Philosophy. Ithaca: Cornell University Press, (1967).
 106. Wallace, J.: "On the Frame of Reference", in [26].
 107. ---- "Nonstandard Theories of Truth", in [27].
 108. ---- "Logical Form, Meaning, Translation", in [47].
 109. Wiggins, David: "The De Re 'Must': a Note on the Logical Form of Essentialist Claims", in [40].
 110. ---- Sameness and Substance. Oxford: Blackwell, (1980).
 111. ---- "'Most' and 'All': Some Comments on a Familiar Programme, and on the Logical Form of Quantified Sentences", in [82].
 112. ---- "Contingency, Identity, and de re and de dicto Necessity", in Dancy, J. (ed.), Papers on Language and Logic. Published by Keele University, (1980).
 113. Wright, Crispin: "Truth Conditions and Criteria", PASSV L (1976), pp. 217-45.

Addenda

114. Lehrer, K. and Taylor, R.: "Time, Truth and Modalities"
Mind LXXIV (1965).
115. Niemi, G.: "On the Existence of a Modal Antinomy",
Synthese 23 (1972), pp. 463-76.
116. Partee, Barbara H.: "The Semantics of Belief Sentences"
in Hintikka, K.J.J. et al. (eds.): Approaches to
Natural Language. Dordrecht: Reidel, (1973).